N. I. AKHIEZER

THEORY OF APPROXIMATION

$$|I_1| \leqq 2A\varphi\left(\frac{1}{\lambda}\right),$$

$$|I_2| \leqq \frac{2B}{\lambda}\int_{\frac{1}{\lambda}}^{\delta}\left|\frac{f(x+t)+f(x-t)}{2}-s\right|\frac{dt}{t^2}$$

$$\leqq \frac{2B}{\lambda}\frac{\varphi(t)}{t}\Big|_{t=\frac{1}{\lambda}}^{\delta}+\frac{4B}{\lambda}\int_{\frac{1}{\lambda}}^{\delta}\frac{\varphi(t)}{t^2}\,dt \leqq$$

$$f(x'')-f(x') = f(\xi+h)-f(\xi-h) =$$

$$= -2\sum_{k=1}^{\infty} a^k \sin(b^k\xi)\sin(b^kh) =$$

$$= -2\sum_{k=1}^{n-1} a^k \sin(b^k\xi)\sin(b^kh),$$

$$\sin(b^kh) = \sin(\pi b^{k-n}) = 0$$

$$f(x'')-f(x') = R+S$$

$$R = -2\sum_{k=1}^{r-1} a^k \sin(b^k\xi)\sin(b^kh),$$

$$S = -2\sum_{k=r}^{n-1} a^k \sin(b^k\xi)\sin(b^kh),$$

$$S = -2\sum_{k=r}^{n-1} a^k \sin(b^kh).$$

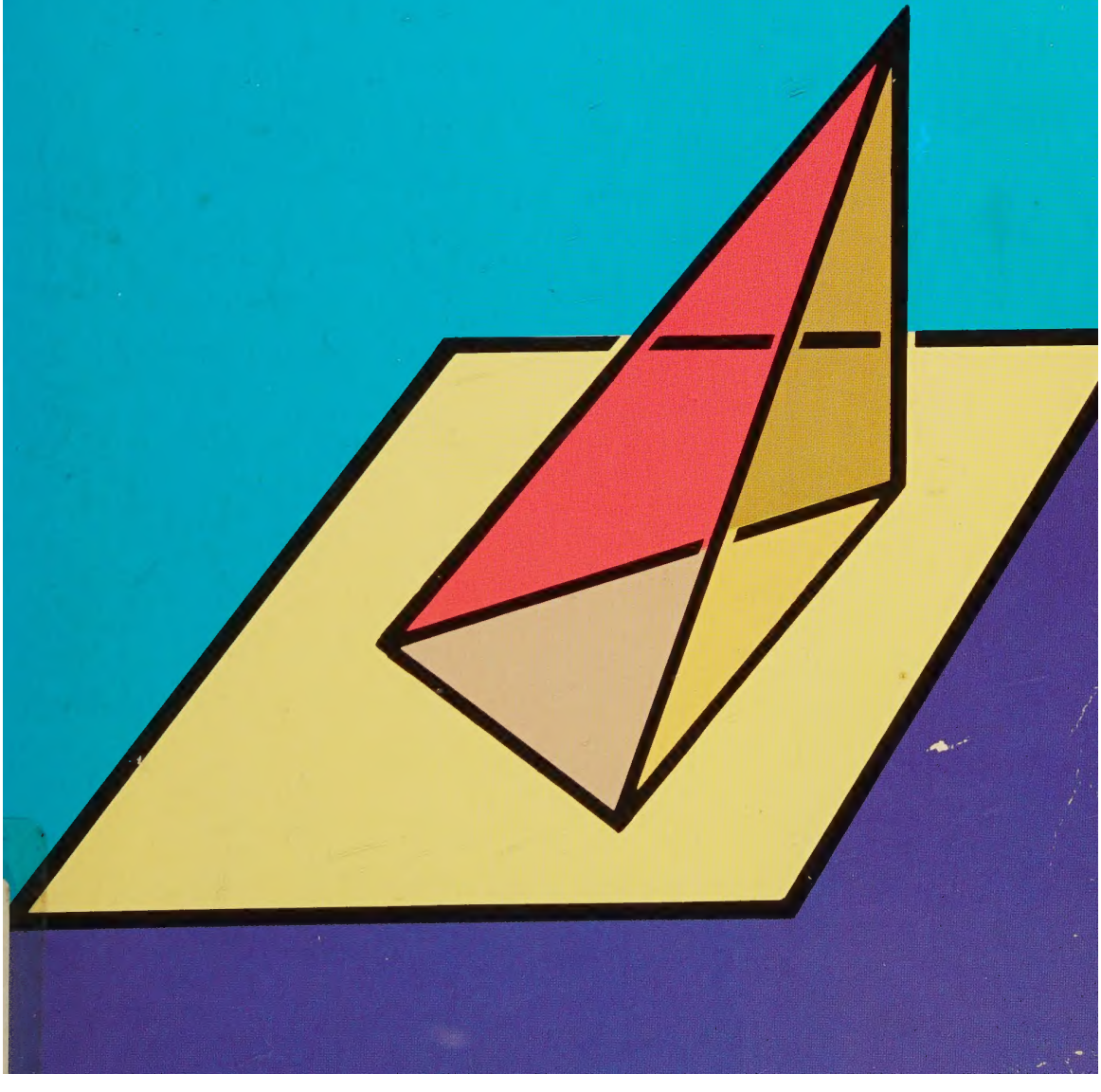
Operations of approximation, whereby functions on sets of a general space are approximated by special functions on such sets, and where the degree of approximation can be judged by defined criteria so that we can speak of an optimal approximation, are found throughout the domain of analysis in its different branches. Using the most modern approach, the author has combined the numerous contributions of analysis to this circle of problems into one coherent theory.

Starting with problems of approximation in Banach and specially in Hilbert spaces, the text proceeds to the theorems of Weierstrass on the approximation of continuous and periodically continuous functions by algebraic and trigonometric polynomials. The Tchebysheff theorems on approximation by rational functions undergo thorough treatment. A long chapter is devoted to the theory of trigonometric series and Fourier integrals in their relation to approximation problems. The modern developments are characterized by names like Akhiezer, Bernstein, Carleman, Fejér, Krein, and many others.

The present book, written by Akhiezer in extraordinarily clear form, represents a brilliant contribution to the treatment of special problems of analysis from the viewpoint of modern functional analysis and is invaluable to mathematicians and physicists.

N.I. Achieser

THEORY OF APPROXIMATION



THEORY OF APPROXIMATION

N. I. ACHIESER

THEORY OF APPROXIMATION

Translated by
CHARLES J. HYMAN

FREDERICK UNGAR PUBLISHING CO.
NEW YORK

Original Russian edition:
LEKTSII PO TEORII APROKSIMATSII

Copyright 1956 by Frederick Ungar Publishing Co.

Printed in the United States of America

Library of Congress Catalog Card Number 56-11459

TRANSLATOR'S PREFACE

This book is the outgrowth of the author's lectures on approximation theory, given at the University of Kharkov.

The subject matter is treated from the viewpoint of functional analysis (or, as we might say, from a higher standpoint). With the modern trend in both pure and applied mathematics toward generalization to abstract spaces (especially Banach and Hilbert spaces), the presentation here of criteria for the "best" approximation of functions in these spaces by means of more convenient functions is of importance to the mathematician and the physicist.

A great deal of the material is taken from original papers, and is not to be found elsewhere; some theorems have never before been published. But even the classical topics, such as Weierstrass' theorems on the approximation of continuous functions by polynomials and of periodic continuous functions by trigonometric sums, have been treated in an interesting fashion and brought into harmony with the general tenor of the whole work.

The reader will welcome Prof. Achieser's tendency to generalize theorems, as for example Hadamard's theorem on determinants, Tchebysheff's approximation by rational functions, de la Vallée-Poussin's approximation theorem and S. N. Bernstein's theorems.

Among the many other features, we mention in particular the approximation of the Weierstrass nondifferentiable function, the approximate solution of an infinite system of linear equations, and the criterion for the representation of a continuous function as a Fourier-Stieltjes integral.

The modern developments in our subject are due chiefly to Achieser, S. N. Bernstein, Favard, Fejér, Haar, Jackson, Krein, F. Riesz and Wiener.

New York, 1956 CHARLES J. HYMAN

CONTENTS

CHAPTER I

APPROXIMATION PROBLEMS IN LINEAR NORMALIZED SPACES

		Page
1.	Formulation of the Principal Problem in the Theory of Approximation	1
2.	The Concept of Metric Space	1
3.	The Concept of Linear Normalized Space	2
4.	Examples of Linear Normalized Spaces	3
5.	The Inequalities of Hölder and Minkowski	4
6.	Additional Examples of Linear Normalized Spaces	7
7.	Hilbert Space	8
8.	The Fundamental Theorem of Approximation Theory in Linear Normalized Spaces	10
9.	Strictly Normalized Spaces	11
10.	An Example of Approximation in the Space L^p	12
11.	Geometric Interpretation	13
12.	Separable and Complete Spaces	14
13.	Approximation Theorems in Hilbert Space	15
14.	An Example of Approximation in Hilbert Space	19
15.	More About the Approximation Problem in Hilbert Space	21
16.	Orthonormalized Vector Systems in Hilbert Space	22
17.	Orthogonalization of Vector Systems	23
18.	Infinite Orthonormalized Systems	25
19.	An Example of a Non-Separable System	29
20.	Weierstrass' First Theorem	29
21.	Weierstrass' Second Theorem	32
22.	The Separability of the Space C	33
23.	The Separability of the Space L^p	34
24.	Generalization of Weierstrass' Theorem to the Space L^p	37
25.	The Completeness of the Space L^p	38
26.	Examples of Complete Orthonormalized Systems in L^2	40
27.	Müntz's Theorem	43
28.	The Concept of the Linear Functional	46
29.	F. Riesz's Theorem	47
30.	A Criterion for the Closure of a Set of Vectors in Linear Normalized Spaces	49

Chapter II

P. L. TCHEBYSHEFF'S DOMAIN OF IDEAS

		Page
31.	Statement of the Problem	51
32.	A Generalization of the Theorem of de la Vallée-Poussin	52
33.	The Existence Theorem	53
34.	Tchebysheff's Theorem	55
35.	A Special Case of Tchebysheff's Theorem	57
36.	The Tchebysheff Polynomials of Least Deviation from Zero	57
37.	A Further Example of P. Tchebysheff's Theorem	58
38.	An Example for the Application of the General Theorem of de la Vallée-Poussin	60
39.	An Example for the Application of P. L. Tchebysheff's General Theorem	62
40.	The Passage to Periodic Functions	64
41.	An Example of Approximating with the Aid of Periodic Functions	66
42.	The Weierstrass Function	66
43.	Haar's Problem	67
44.	Proof of the Necessity of Haar's Condition	68
45.	Proof of the Sufficiency of Haar's Condition	69
46.	An Example Related to Haar's Problem	72
47.	P. L. Tchebysheff's Systems of Functions	73
48.	Generalization of P. L. Tchebysheff's Theorem	74
49.	On a Question Pertaining to the Approximation of a Continuous Function in the Space L	76
50.	A. A. Markoff's Theorem	82
51.	Special Cases of the Theorem of A. A. Markoff	85

Chapter III

ELEMENTS OF HARMONIC ANALYSIS

52.	The Simplest Properties of Fourier Series	89
53.	Fourier Series for Functions of Bounded Variation	93
54.	The Parseval Equation for Fourier Series	97
55.	Examples of Fourier Series	98
56.	Trigonometric Integrals	101
57.	The Riemann-Lebesgue Theorem	103
58.	Plancherel's Theory	104
59.	Watson's Theorem	106
60.	Plancherel's Theorem	108
61.	Fejér's Theorem	110
62.	Integral-Operators of the Fejér Type	113
63.	The Theorem of Young and Hardy	116
64.	Examples of Kernels of the Fejér Type	118
65.	The Fourier Transformation of Integrable Functions	120

Page
66. The Faltung of two Functions 122
67. V. A. Stekloff's Functions 123
68. Multimonotonic Functions 125
69. Conjugate Functions 126

Chapter IV

CERTAIN EXTREMAL PROPERTIES OF INTEGRAL TRANSCENDENTAL FUNCTIONS OF THE EXPONENTIAL TYPE

70. Integral Functions of the Exponential Type 130
71. The Borel Transformation 132
72. The Theorem of Wiener and Paley 134
73. Integral Functions of the Exponential Type which are Bounded along the Real Axis 137
74. S. N. Bernstein's Inequality 140
75. B. M. Levitan's Polynomials 146
76. The Theorem of Fejér and Riesz. A Generalization of This Theorem 152
77. A Criterion for the Representation of Continuous Functions as Fourier-Stieltjes Integrals 154

Chapter V

QUESTIONS REGARDING THE BEST HARMONIC APPROXIMATION OF FUNCTIONS

78. Preliminary Remarks 160
79. The Modulus of Continuity. 161
80. The Generalization to the Space L^p $(p \geq 1)$ 162
81. An Example of Harmonic Approximation 165
82. Some Estimates for Fourier Coefficients 169
83. More about V. A. Stekloff's Functions. 173
84. Two Lemmas. 175
85. The Direct Problem of Harmonic Approximation 176
86. A Criterion due to B. Sz.-Nagy 183
87. The Best Approximation of Differentiable Functions 187
88. Direct Observations Concerning Periodic Functions 195
89. Jackson's Second Theorem. 199
90. The Generalized Fejér Method 201
91. Berstein's Theorem 206
92. Priwaloff's Theorem 210
93. Generalizations of Bernstein's Theorems to the Space L^p $(p \geq 1)$ . . 211
94. The Best Harmonic Approximation of Analytic Functions 214
95. A Different Formulation of the Result of the Preceding Section. . . 218
96. The Converse of Bernstein's Theorem 221

Chapter VI

WIENER'S THEOREM ON APPROXIMATION

Page

97. Wiener's Problem 224
98. The Necessity of Wiener's Condition 224
99. Some Definitions and Notation 225
100. Several Lemmas 227
101. The Wiener-Levy Theorem. 230
102. Proof of the Sufficiency of Wiener's Condition 233
103. Wiener's General Tauber Theorem 234
104. Weakly Decreasing Functions 235
105. Remarks on the Terminology. 237
106. Ikehara's Theorem 238
107. Carleman's Tauber Theorem 241

VARIOUS ADDENDA AND PROBLEMS

A. Elementary Extremal Problems and Certain Closure Criteria 243
B. Szegö's Theorem and Some of Its Applications 256
C. Further Examples of Closed Sequences of Functions 267
D. The Carathéodory-Fejér Problem and Similar Problems 270
E. Solotareff's Problems and Related Problems 280
F. The Best Harmonic Approximation of the Simplest Analytic Functions 289

Notes 296
Index 306

CHAPTER I

APPROXIMATION PROBLEMS IN LINEAR NORMALIZED SPACES

1. Formulation of the Principal Problem in the Theory of Approximation. The main problem in the theory of approximation can be stated as follows: Let us suppose that two functions $f(P)$ and $F(P; A_1, \ldots, A_n)$ of the point $P \in \mathfrak{P}$ are defined within a certain point set $\mathfrak{P}$ in a space of any number of dimensions. Here $F(P; A_1, \ldots, A_n)$ depends on a certain number of parameters A_r. It is required to so determine the parameters that the deviation of the function $F(P; A_1, \ldots, A_n)$ from the function $f(P)$ in $\mathfrak{P}$ shall be a minimum. Of course, we must define what we mean by the deviation or, as we sometimes say, the distance between the function F and the function f.

If we consider, for example, bounded functions, we can take as the distance between two functions the least upper bound of the absolute value of their difference within $\mathfrak{P}$. This definition of distance brings to light many relations which are valid, not only for points in ordinary space, but also for the totality of all bounded functions within $\mathfrak{P}$.

This latter circumstance, which is encountered regularly in mathematics in the study of other classes of functions and many other totalities (sets), has led to the creation of the extremely important concept of metric space.

2. The Concept of Metric Space. *A set E having the elements $x, y, z, \ldots$ is known as metric space* [1], [1] *and the elements are called points of the space*, if for every pair of elements x, y there can be found a corresponding non-negative number $D[x, y]$, which is called the distance between the points x and y, and which satisfies the following conditions:

A. $D[x, x] = 0$,
B. $D[x, y] = D[y, x] > 0$ (if $x \neq y$),
C. $D[x, z] \leq D[x, y] + D[y, z]$ (triangular inequality).

E.g. the set of all sequences of the (in general, complex) numbers

$$x = \{x_1, x_2, \ldots\} = \{x_k\}, \quad y = \{y_1, y_2, \ldots\} = \{y_k\}, \ldots$$

[1] The numbers in brackets refer to the notes at the end of the book.

forms a metric space if we presuppose that $x = y$ if and only if $x_k = y_k$ ($k = 1, 2, \ldots$) and the distance is defined by means of the formula

$$D[x, y] = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot \frac{|x_k - y_k|}{1 + |x_k - y_k|} .$$

From the inequality

$$\frac{a+b}{1+a+b} \leqq \frac{a}{1+a} + \frac{b}{1+b} \quad (a \geqq 0, b \geqq 0),$$

which is easily proved, it follows that the condition C is fulfilled.

3. The Concept of Linear Normalized Space [1]. Among metric spaces, the so-called linear normalized spaces are particularly important.

A set E having the elements $x, y, z, \ldots$ is called linear normalized space, the elements themselves points or vectors, if

1. there is defined in E an operation, which we call addition and denote by the symbol $+$, in respect to which E forms an Abelian group; the zero element of the group E will be denoted by 0;

2. a multiplication of the elements of the set E by (real or complex) numbers $\alpha, \beta, \gamma, \ldots$ is defined so that

$$\begin{aligned} \alpha(x+y) &= \alpha x + \alpha y, \\ (\alpha+\beta)x &= \alpha x + \beta x, \\ \alpha(\beta x) &= (\alpha\beta)x, \\ 1 \cdot x &= x, \\ 0 \cdot x &= 0 ; \end{aligned}$$

3. to every element $x \in E$ there corresponds a certain positive number $\| x \|$ called the *norm* of the element x, which satisfies the conditions

$\| x \| = 0$ if and only if $x = 0$,
$\| \alpha x \| = |\alpha| \cdot \| x \|$,
$\| x + y \| \leq \| x \| + \| y \|$.

n vectors $x_1, x_2, \ldots, x_n$ are said to be *linearly independent* if the equation

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0$$

entails the equations

$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0 .$$

A linear normalized space becomes a metric space if we put $D[x, y] = \| x - y \|$ with $x - y = x + (-1)y$.

4. Examples of Linear Normalized Spaces. [2] *The space C.* The totality of all continuous functions $x = x(P)$ of the points P of a bounded closed set $\mathfrak{P}$ in ordinary space of any dimensionality is a linear normalized space (the space C) if we put $x = y$ for $x(P) \equiv y(P)$. Furthermore, addition and multiplication by a number shall be understood in the usual sense and the norm defined by the equation

$$\| x \| = \| x \|_C = \max_{P \in \mathfrak{P}} | x(P) | .$$

The case in which $\mathfrak{P}$ is a finite interval of the number axis has special significance. Since we can pass from the interval [3] $[a, b]$ to an arbitrary closed interval by means of a substitution of the form

$$t = At' + B$$

we usually take the interval $[0, 1]$ or $[-1, 1]$.

On the other hand, in the cases involving only periodic functions, we ordinarily take the interval $[-\pi, \pi]$ or any other of length 2π and count its ends as one point.

By means of the substitution

$$u = \operatorname{tg} \frac{t}{2} \qquad (-\pi \leqq t \leqq \pi)$$

we can pass from the space of continuous periodic functions to the space of those functions $x = x(u)$ which are continuous along the entire axis and whose limits

$$\lim_{u \to -\infty} x(u), \qquad \lim_{u \to \infty} x(u)$$

exist and are equal to each other.

We will denote this space by C_∞.

In considering C_∞, we may count the points $\pm \infty$ as one point. In like manner, by means of appropriate substitutions of variables, we can pass from the space $C(0, 1)$ of all (not merely the periodic) functions which are continuous in the interval $[0, 1]$ to the spaces $C(0, \infty)$ and $C(-\infty, \infty)$ of functions which are continuous in the closed intervals $[0, \infty]$ and $[-\infty, \infty]$.

The space $C(-\infty, \infty)$ is the totality of all continuous functions $x(t)$ $(-\infty < t < \infty)$ such that the limits

$$\lim_{t \to -\infty} x(t), \qquad \lim_{t \to \infty} x(t)$$

exist. However, these limits might be unequal. This fact distinguishes $C(-\infty, \infty)$ from C_∞.

---

[2] Unless otherwise stated, all numbers, functions and sequences are complex.

[3] A closed interval, i.e. an interval inclusive of its ends, will always be denoted by letters within brackets.

The space L^p $(p \geqq 1)$. By L^p is meant the totality of all functions measurable in the (finite or infinite) interval[4] (a, b), whose absolute value to the p—th power is integrable in the sense of Lebesgue. In this connection, addition and multiplication with complex numbers are to be considered in the ordinary sense. Two elements $x = x(t)$, $y = y(t) \in L^p$ are identified if the equality $x(t) = y(t)$ holds almost everywhere.

The norm is defined by the equation

$$\| x \| = \| x \|_p = \left\{ \int_a^b | x(t) |^p dt \right\}^{\frac{1}{p}}.$$

If it proves necessary to indicate the interval, we employ the notation $L^p(a, b)$ instead of L^p. In place of L^1 we write simply L.

We will show that L^p is a linear normalized space. Since the sum of two measurable functions is a measurable function and the inequality[5]

$$| \alpha + \beta |^p \leqq 2^p (| \alpha |^p + | \beta |^p)$$

obviously holds true, it follows from $x, y \in L^p$ that $x + y \in L^p$.

A function which is zero almost everywhere in the interval is called the zero element.

If $x \in L^p$ and α is any number, then $\alpha x \in L^p$.

Hence it remains to prove the inequality

$$\left\{ \int_a^b | x(t) + y(t) |^p dt \right\}^{\frac{1}{p}} \leqq \left\{ \int_a^b | x(t) |^p dt \right\}^{\frac{1}{p}} + \left\{ \int_a^b | y(t) |^p dt \right\}^{\frac{1}{p}},$$

which is known as *Minkowski's Inequality.*

5. The Inequalities of Hölder and Minkowski [2]. Suppose that $\alpha, \beta \geqq 0$, $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ (so that $q > 1$).

Let us study the curve OAB, whose equation is

$$y = x^{p-1}$$

---

[4] Instead of an interval, we can take as the defined region of the function any measurable set (of finite or infinite measure) on the number axis or in ordinary space of any dimensionality.

[5] It will be observed that also the inequality

$$| \alpha + \beta |^p \leqq 2^{p-1} (| \alpha |^p + | \beta |^p) \qquad (p \geqq 1)$$

is valid; for the verification thereof, it suffices to find the minimum of $A^p + B^p$, subject to the conditions $A + B = 1$, $A \geqq 0$, $B \geqq 0$.

or, what amounts to the same thing,

$$x = y^{q-1}.$$

From Fig. 1, [6] it is evident that the sum of the areas S and T is not less than $\alpha\beta$, and the equality obtains if and only if the points A and B coincide, i.e. if $\alpha^p = \beta^q$. Whence it follows that

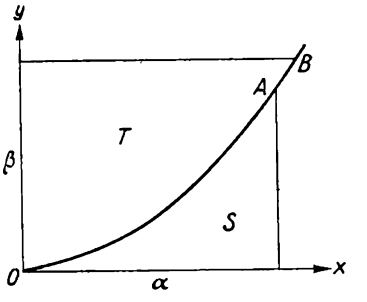


Fig. 1

$$\alpha\beta \leqq \frac{\alpha^p}{p} + \frac{\beta^q}{q}. \tag{1}$$

When $p = 2$ (and hence $q = 2$), the inequality (1) becomes the elementary inequality

$$2\alpha\beta \leqq \alpha^2 + \beta^2 .$$

We will assume that $u(t) \in L^p$, $v(t) \in L^q$, also that

$$\alpha = \frac{|u(t)|}{||u||_p}, \qquad \beta = \frac{|v(t)|}{||v||_q}$$

and utilize the inequality (1). We then have

$$\frac{|u(t)v(t)|}{||u||_p ||v||_q} \leqq \frac{1}{p}\frac{|u(t)|^p}{\{||u||_p\}^p} + \frac{1}{q}\frac{|v(t)|^q}{\{||v||_q\}^q}.$$

As the right-hand side is integrable, so is the left-hand side, and the integral of the expression on the left is not greater than the integral of the expression on the right, i.e. not greater than

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Thus we have found Hölder's Inequality

$$\int_a^b |u(t)v(t)|\,dt \leqq \left\{\int_a^b |u(t)|^p\,dt\right\}^{\frac{1}{p}} \cdot \left\{\int_a^b |v(t)|^q\,dt\right\}^{\frac{1}{q}}. \tag{2}$$

As can be readily seen, the equality sign in this relation is valid if and only if, almost everywhere in (a, b), we have

$$|v(t)|^q = C\,|u(t)|^p .$$

---

[6] This corresponds to the case in which $p > 2$ and $\beta > \alpha^{p-1}$.

In the case $p = q = 2$, equation (2) is usually called *Schwarz's Inequality*, although strictly speaking it should be called Buniakovski's Inequality. [7]

To obtain Minkowski's inequality, we observe: From

$$| x(t) | + | y(t) | \in L^p \quad . (p > 1)$$

it follows that

$$\{ | x(t) | + | y(t) | \}^{p-1} \in L^q ,$$

if we remember that

$$\frac{1}{p} + \frac{1}{q} = 1 \cdot$$

Applying Hölder's inequality to every term on the right-hand side of the equation

$$(| x | + | y |)^p = | x | \cdot (| x | + | y |)^{p-1} + | y | \cdot (| x | + | y |)^{p-1}$$

we obtain

$$\int_a^b (| x | + | y |)^p \, dt \leqq$$

$$\leqq \left\{ \int_a^b (| x | + | y |)^p \, dt \right\}^{\frac{1}{q}} \left\{ \left[\int_a^b | x |^p \, dt \right]^{\frac{1}{p}} + \left[\int_a^b | y |^p \, dt \right]^{\frac{1}{p}} \right\} .$$

Whence Minkowski's inequality follows when $p > 1$. When $p = 1$, this inequality is obvious.

It is easy to see that in Minkowski's inequality (i.e. when $p > 1$) the equality sign is valid if and only if $x(t) = Cy(t)$ almost everywhere, C being a constant $\geqq 0$.

It will be noted that this method yields also the Hölder and Minkowski inequalities for series. We have

$$\sum_{k=1}^{\infty} | \alpha_k \beta_k | \leqq \left\{ \sum_{k=1}^{\infty} | \alpha_k |^p \right\}^{\frac{1}{p}} \left\{ \sum_{k=1}^{\infty} | \beta_k |^q \right\}^{\frac{1}{q}} \quad \left(\frac{1}{p} + \frac{1}{q} = 1, \quad p > 1 \right),$$

$$\left\{ \sum_{k=1}^{\infty} | \alpha_k + \beta_k |^p \right\}^{\frac{1}{p}} \leqq \left\{ \sum_{k=1}^{\infty} | \alpha_k |^p \right\}^{\frac{1}{p}} + \left\{ \sum_{k=1}^{\infty} | \beta_k |^p \right\}^{\frac{1}{p}} \quad (p \geqq 1).$$

Moreover, it is not difficult to see that these inequalities are special cases of the Hölder and Minkowski inequalities for integrals, established above.

---

[7] Translator's note: This inequality was published by H. A. Schwarz in 1885. He used it as an aid in the solution of an extremum problem in surface theory. The work of V. Buniakovski had already been published in 1859. V. Buniakovski, *Sur quelques inégalités concernant les intégrales ordinaires et les intégrales aux différences finies.* Mémoires de l'Académie des Sciences de St. Pétersbourg, 7th Series (1859).

In what follows, we shall have occasion to make use of the so-called generalized Minkowski inequality for integrals, which has the form

$$\left\{\int_a^b \left|\int_c^d \varphi(x, y)\, dy\right|^p dx\right\}^{\frac{1}{p}} \leqq \int_c^d \left\{\int_a^b |\varphi(x, y)|^p dx\right\}^{\frac{1}{p}} dy \quad (p \geqq 1)$$

This inequality is obtained [3] in the theory of integrals by the usual passage from a function which assumes only a finite number of values (at least two values) to an arbitrary function $\varphi(x, y)$, measurable in a rectangular region R ($a < x < b$, $c < y < d$), such that the right-hand side has meaning.

As a special case thereof, if $p = 1$, we have the theorem of Fubini [3]. This asserts: From

$$\int_c^d \left\{\int_a^b |\varphi(x, y)|\, dx\right\} dy < \infty$$

it follows that

$$\int_R \varphi(P)\, dP = \int_a^b \left\{\int_c^d \varphi(x, y)\, dy\right\} dx = \int_c^d \left\{\int_a^b \varphi(x, y)\, dx\right\} dy.$$

6. Additional Examples of Linear Normalized Spaces. *The spaces m and c.* The set m of all bounded number sequences and the set c of all convergent number sequences, viz,

$$x = \{x_k\}, \quad y = \{y_k\}, \ldots$$

form linear normalized spaces if we postulate that

$$x + y = \{x_k + y_k\},$$
$$ax = \{ax_k\},$$
$$\| x \| = \sup |x_k|$$

The space l^p ($p \geq 1$). This denotes the set of all sequences

$$x = \{x_k\}, \quad y = \{y_k\}, \ldots,$$

for which the series

$$\sum_{k=1}^{\infty} |x_k|^p, \quad \sum_{k=1}^{\infty} |y_k|^p, \ldots$$

converge. l^p will become a linear normalized space if we assume

$$\| x \| = \left\{\sum_{k=1}^{\infty} |x_k|^p\right\}^{\frac{1}{p}}$$

The reader will satisfy himself on that score by considering Minkowski's inequality for series.

7. Hilbert Space [4]. We define Hilbert space (space H) as a linear space (i.e. the first two conditions of Section 3 are fulfilled) in which for every pair of vectors x and y there is a corresponding (complex) number (x, y), called the *scalar product* of the vectors x and y, and satisfying the following conditions [8]:

a) $$(y, x) = \overline{(x, y)},$$

b) $$(a_1 x_1 + a_2 x_2, y) = a_1 (x_1, y) + a_2 (x_2, y),$$

c) $$(x, x) \geqq 0,$$

d) $(x, x) = 0$ if and only if $x = 0$.

The vectors x and y are called *orthogonal* if $(x, y) = 0$.

By giving concrete examples, it is easy to show that we can realize the conditions by means of which we have defined Hilbert's space axiomatically (it is therefore natural that H be designated as abstract Hilbert space).

The space l^2, for example, furnishes one such concrete realization if, by hypothesis,

$$(x, y) = \sum_{k=1}^{\infty} x_k \overline{y}_k \ ,$$

where

$$x = \{x_k\}, \quad y = \{y_k\}.$$

The convergence (even the absolute) of the series

$$\sum_{k=1}^{\infty} x_k \overline{y}_k$$

follows from Schwarz's inequality for series (already known to Cauchy), that is, from Hölder's inequality for $p = q = 2$.

The space L^2 represents another example if we put

$$(x, y) = \int_a^b x(t)\, \overline{y(t)}\, dt \ .$$

We will show that the scalar product (x, y) has the following properties:

1°. $$(ax, ax) = |a|^2 (x, x),$$

[8] A horizontal bar over a number denotes the passage to the conjugate complex number.

2^0. $$|(x, y)| \leqq \sqrt{(x, x)} \cdot \sqrt{(y, y)}$$

(generalized Schwarzian inequality),

3^0. $$\sqrt{(x+y, x+y)} \leqq \sqrt{(x, x)} + \sqrt{(y, y)}.$$

In the foregoing, if $y \neq 0$, the equality sign is valid in the relation 2° only when $x = ay$ and in the relation 3° only when $x = ay$ and $a \geq 0$. (If $y = 0$, we get a trivial equation).

The property 1° is obvious:

$$(ax, ax) = a(x, ax) = a\overline{(ax, x)} = a\bar{a}\overline{(x, x)} = a\bar{a}(x, x).$$

To establish the property 2°, we need to assume that $(x, y) \neq 0$. By setting

$$\vartheta = \frac{(x, y)}{|(x, y)|},$$

we find that, for any real λ,

$$0 \leqq (\bar{\vartheta}x + \lambda y, \bar{\vartheta}x + \lambda y) = \lambda^2 (y, y) + 2\lambda |(x, y)| + (x, x)$$

the roots of this polynomial must be either complex or equal to each other, hence the inequality 2° is established. The equality sign holds if and only if, for a given λ,

$$\bar{\vartheta}x + \lambda y = 0 \cdot$$

We will now establish the property 3°. We have

$$(x+y, x+y) = (x, x) + (x, y) + (y, x) + (y, y) \leqq (x, x) + 2|(x, y)| + (y, y),$$

and by virtue of 2° we obtain

$$(x+y, x+y) \leqq \left\{\sqrt{(x, x)} + \sqrt{(y, y)}\right\}^2.$$

The equality sign holds if and only if

$$(x, y) = \sqrt{(x, x)} \cdot \sqrt{(y, y)} \ .$$

As we know, if $y \neq 0$, it is necessary that $x = ay$. If this condition is fulfilled, we have

$$(x, y) = (ay, y) = a(y, y),$$

and consequently

$$a = \sqrt{\frac{(x, x)}{(y, y)}} \geqq 0.$$

As a result of the conditions defining the scalar product and of properties 1°, 3° of said scalar product, Hilbert's space represents a linear normalized space if we put

$$\| x \| = \sqrt{(x, x)}.$$

8. The Fundamental Theorem of Approximation Theory in Linear Normalized Spaces. Let E be any linear normalized space, and let

$$g_1, g_2, \ldots, g_n$$

be n linearly independent elements of E. The leading problem in approximations for the "linear case," now being considered, can be formulated as follows: Given an element $x \in E$; it is required to determine numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that the quantity

$$\varphi(\lambda_1, \lambda_2, \ldots, \lambda_n) = \| x - \lambda_1 g_1 - \lambda_2 g_2 - \ldots - \lambda_n g_n \|$$

shall attain its smallest value.

In this section, we will prove that numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ having the required property do exist.

First of all, let us observe that $\varphi(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a continuous function of its arguments. In fact, by virtue of the triangular inequality, we have:

$$| \varphi(\lambda_1', \lambda_2', \ldots, \lambda_n') - \varphi(\lambda_1, \lambda_2, \ldots, \lambda_n) | =$$

$$= \left| \left\| x - \sum_1^n \lambda_k' g_k \right\| - \left\| x - \sum_1^n \lambda_k g_k \right\| \right| \leqq \left\| \sum_1^n (\lambda_k' - \lambda_k) g_k \right\| \leqq$$

$$\leqq \sum_1^n | \lambda_k' - \lambda_k | \, \| g_k \| \leqq \max_k | \lambda_k' - \lambda_k | \cdot \sum_1^n \| g_r \|.$$

We now introduce a second continuous function

$$\psi(\lambda_1, \lambda_2, \ldots, \lambda_n) = \| \lambda_1 g_1 + \lambda_2 g_2 + \ldots + \lambda_n g_n \| .$$

According to a well-known theorem of Weierstrass, the function ψ has a minimum μ upon the "sphere"

$$| \lambda_1 |^2 + | \lambda_2 |^2 + \ldots + | \lambda_n |^2 = 1,$$

which is a bounded closed point set in ordinary space of a finite number of dimensions.

The non-negative number μ cannot be zero, since the vectors $g_1, \ldots, g_n$ are linearly independent. Hence $\mu > 0$.

Let ϱ ($\geqq 0$) denote the greatest lower bound of the values of the function φ. If

$$\sqrt{\sum_1^n | \lambda_k |^2} > \frac{1}{\mu} (\varrho + 1 + \| x \|) = R$$

it follows that

$$\varphi(\lambda_1, \lambda_2, \ldots, \lambda_n) \geqq \| \lambda_1 g_1 + \ldots + \lambda_n g_n \| - \| x \| \geqq$$
$$\geqq \sqrt{\sum_1^n |\lambda_k|^2} \cdot \mu - \| x \| > \varrho + 1.$$

Therefore, in seeking the minimum of the function $\varphi(\lambda_1, \lambda_2, \ldots, \lambda_n)$ we can restrict ourselves to those values of λ_k for which

$$\sum_1^n |\lambda_k|^2 \leqq R^2,$$

in other words, to the consideration of the function φ in a bounded closed region. In such a region, a continuous function has a minimum.

Accordingly, the existence of the linear expression $\varphi(\lambda_1, \lambda_2, \ldots, \lambda_n)$, which furnishes the best approximation of the element x, has been demonstrated.

9. Strictly Normalized Spaces. The question arises: When is the expression $\lambda_1 g_1 + \lambda_2 g_2 + \ldots + \lambda_n g_n$, which furnishes the best approximation of the element x, uniquely determined for all values of $x \in E$?

The aforesaid uniqueness obtains whenever the space is strictly normalized, [9] i.e. if the equality sign in the inequality

$$\| x + y \| \leqq \| x \| + \| y \| \qquad (x \neq 0, y \neq 0)$$

holds only for $y = \alpha x$ $(\alpha \geq 0)$.

Let us suppose that the space E is strictly normalized, and assume that two expressions

$$\lambda_1 g_1 + \lambda_2 g_2 + \ldots + \lambda_n g_n, \qquad \mu_1 g_1 + \mu_2 g_2 + \ldots \mu_n g_n$$

(in which $g_1, g_2, \ldots, g_n$ are linearly independent) yield the best approximation for the element x. Since, by hypothesis (see Section 8),

$$\left\| x - \sum_1^n \lambda_k g_k \right\| = \left\| x - \sum_1^n \mu_k g_k \right\| = \varrho$$

we can evidently assume that $\varrho \neq 0$.

It follows from

$$\left\| x - \sum_1^n \frac{\lambda_k + \mu_k}{2} g_k \right\| \leqq \frac{1}{2} \left\| x - \sum_1^n \lambda_k g_k \right\| + \frac{1}{2} \left\| x - \sum_1^n \mu_k g_k \right\| = \varrho$$

that

$$\left\| x - \sum_1^n \frac{\lambda_k + \mu_k}{2} g_k \right\| = \varrho$$

[9] This expression is due to M. Krein.

and

$$\left\| x - \sum_1^n \frac{\lambda_k + \mu_k}{2} g_k \right\| = \frac{1}{2} \left\| x - \sum_1^n \lambda_k g_k \right\| + \frac{1}{2} \left\| x - \sum_1^n \mu_k g_k \right\|.$$

Hence, because the space is strictly normalized, we have

$$x - \sum_1^n \lambda_k g_k = \alpha \left\{ x - \sum_1^n \mu_k g_k \right\}.$$

In this relation, α can only equal 1, because otherwise the element x would be a linear combination of the elements $g_1, \ldots, g_n$, and hence ϱ would be zero. But if $\alpha = 1$, then we have

$$\sum_1^n (\lambda_k - \mu_k) g_k = 0$$

and consequently also

$$\lambda_1 = \mu_1, \quad \lambda_2 = \mu_2, \ldots, \quad \lambda_n = \mu_n,$$

since the elements $g_1, g_2, \ldots, g_n$ are linearly independent. Therefore the expressions under consideration are identical.

Space H provides an example of a strictly normalized space, and so does L^p when $p > 1$, while the spaces C and L are not strictly normalized.

In fact, suppose we are given the interval $[-1, 1]$ and two linearly independent functions $x(t), y(t) \in C$ whose moduli assume their maximum values at one and the same point a of the interval and

$$\arg x(a) = \arg y(a);$$

then obviously

$$\| x + y \| = \| x \| + \| y \|.$$

In order to show that L is not a strictly normalized space either, it suffices to put $x(t) = 1$ when $t \geq 0$ and $x(t) = 0$ when $t < 0$ and $y(t) = 1 - x(t)$.

10. An Example of Approximation in the Space L^p [5]. As an example, we present the problem: Among all trigonometric sums of the form

$$F_n(\Theta) = \alpha \cos n\Theta + \beta \sin n\Theta + a_{n-1} \cos (n-1)\Theta + \\ + b_{n-1} \sin (n-1)\Theta + \ldots + a_1 \cos \Theta + b_1 \sin \Theta + a_0$$

where α and β are constants, find those sums for which the integral

$$\int_0^{2\pi} | F_n(\Theta) |^p \, d\Theta$$

attains its smallest value. We will assume at first that $p > 1$ and make use of the fact that, under this hypothesis, the space L^p is strictly normalized. Since the trigonometric functions are periodic,

$$\int_0^{2\pi} |F_n(\Theta)|^p \, d\Theta = \int_{\varphi}^{2\pi+\varphi} |F_n(\Theta)|^p \, d\Theta = \int_0^{2\pi} |F_n(\Theta+\varphi)|^p \, d\Theta$$

for all (real) values of φ. Put $\varphi = \frac{\pi}{2}$ and observe that

$$F_n\left(\Theta + \frac{\pi}{n}\right) = -\alpha \cos n\Theta - \beta \sin n\Theta + a_{n-1} \cos (n-1)\Theta + \\ + b'_{n-1} \sin (n-1)\Theta + \ldots + a'_1 \cos \Theta + b'_1 \sin \Theta + a_0 \, .$$

Because the space is strictly normalized, we arrive at the result that

$$F_n\left(\Theta + \frac{\pi}{n}\right) \equiv -F_n(\Theta)$$

for the required trigonometric sums, and consequently

$$F_n\left(\Theta + \frac{2\pi}{n}\right) \equiv F_n(\Theta) \, .$$

This last relation proves that $F(\Theta)$ has the period $\frac{2\pi}{n}$ and hence

$$a_{n-1} = b_{n-1} = \ldots = a_1 = b_1 = 0 \, ,$$

while the first relation leads to the equation

$$a_0 = 0 \, .$$

Thus, the only solution of our problem for $p > 1$ is the expression

$$F_n(\Theta) = \alpha \cos n\Theta + \beta \sin n\Theta .$$

This expression is also a solution [10] for $p = 1$, as we readily see by recalling the principles of continuity.

11. Geometric Interpretation. The problem for which the existence of a solution was proved in Section 8 permits of a useful geometric interpretation. The totality of all points of the form

$$\alpha_1 g_1 + \alpha_2 g_2 + \ldots + \alpha_n g_n ,$$

where the quantities $\alpha_1, \alpha_2, \ldots, \alpha_n$ run through all the complex numbers for fixed linearly independent elements $g_1, g_2, \ldots, g_n \in E$, represents a certain linear manifold $G \subset E$ of the type such that $g', g'' \in G$ entails $\alpha' g' + \alpha'' g'' \in G$ for any complex quantities α', α''. This linear manifold evidently forms a space, since it contains the point 0. When $n = 1$, we get a "line"; when $n = 2$, a "plane"; and, generally, when n is any integer, an "n-dimensional plane."

[10] And moreover the only solution by virtue of the general theorems in Section 49.

Thus our problem consisted in determining a point in the finite-dimensional subspace G of E such that its distance from the given point $x \in E$ (in reference to the metric of E), is the shortest. In Section 8 we proved that there exists such a point in G.

If E itself is not of finite dimension, i.e. if it has an arbitrary number of linearly independent vectors, then E contains a subspace of infinite dimensionality. Let G be such a subspace.

The question arises whether there exists in G a point whose distance from a given point $x \in E$ is the shortest. In Section 15 this question will be answered in the affirmative, provided that E is a complete Hilbert space. However, at this time we note that in the case of strict normalization of the space E, there can exist in G at most one point at the shortest distance from a given point $x \in E$. The proof thereof is similar to that of the special case considered in Section 9.

12. Separable and Complete Spaces. Let $\mathfrak{M}$ be a point set in the metric space E. A point $x \in E$ is called a *cluster point* of the set $\mathfrak{M}$ if an infinite set of points belonging to $\mathfrak{M}$ is to be found in all "spheres"

$$D[x, x_0] < \varrho$$

(with "center" at x and "radius" ϱ). The point x_0 itself need not be situated in $\mathfrak{M}$. A set $\mathfrak{M}$ is called "closed" if it contains all of its cluster points. If the set $\mathfrak{M}$ is not closed, we can obtain a *closure* of $\mathfrak{M}$ (denoted by the symbol $\overline{\mathfrak{M}}$) by adding to $\mathfrak{M}$ all its cluster points.

The set $\mathfrak{M}$ is said to be "dense" in E if $\overline{\mathfrak{M}} = E$.

Suppose, for example, that E is the totality of all real numbers. Then the totality of all rational numbers is a set which lies dense in E. That follows from the fact that all the real numbers are cluster points of the sequences of rational numbers.

The space E is said to be *separable* if it contains a dense denumerable set of points.

For instance, the aforementioned space of all real numbers is separable. That follows from the fact that the set of rational numbers is denumerable.

We say, the sequence $\{x_n\}$ of points of E *converges towards the point* x, if the relation

$$\lim_{n \to \infty} D[x, x_n] = 0 \tag{1}$$

holds true. That situation will be described by writing

$$x_n \to x .$$

By reason of the triangular inequality, (1) gives immediately

$$\lim_{m, n \to \infty} D[x_m, x_n] = 0 . \tag{2}$$

However, the converse is not necessarily true, that is to say, the relation (2) or, what amounts to the same thing, the convergence by Cauchy's test, does not of necessity imply the existence in E of an element towards which the sequence converges.

The metric space E is said to be *complete* if all sequences therein which converge by Cauchy's test, converge towards a certain element of E.

For example, C is a complete space because convergence in it means uniform convergence, and we know from analysis that a sequence of continuous functions which is uniformly convergent in the Cauchy sense approaches a certain continuous function uniformly.

13. Approximation Theorems in Hilbert Space. Let G be a subspace of the Hilbert space H and suppose that the point $x \in H$ does not belong to G. If there exists in G a point y whose distance from x is the shortest, then the vector $x - y$ is orthogonal to any vector g of G, i.e.

$$(x - y, g) = 0. \qquad (g \in G).$$

To verify this statement, let us assume that there exists in G a vector f such that

$$(x - y, f) = \sigma \neq 0$$

and consider the vector

$$y^* = y + \frac{\sigma}{(f,f)} f \in G.$$

We have

$$\| x - y^* \|^2 = \left(x - y - \frac{\sigma}{(f,f)} f, \quad x - y - \frac{\sigma}{(f,f)} f\right) =$$

$$= \| x - y \|^2 - \frac{\bar{\sigma}}{(f,f)} (x - y, f) - \frac{\sigma}{(f,f)} (f, x - y) + \frac{|\sigma|^2}{(f,f)} = \| x - y \|^2 - \frac{|\sigma|^2}{(f,f)},$$

and hence

$$\| x - y^* \| < \| x - y \|.$$

But this contradicts our hypothesis, to the effect that y is the point of the subspace G which is at the shortest distance from x.

The vector $y \in G$ which has the property that the difference $x - y$ is orthogonal to G is naturally called the projection of x upon G.

By employing a proven theorem, we can present explicitly the vector

$$y = \lambda_1 g_1 + \lambda_2 g_2 + \ldots + \lambda_n g_n,$$

which deviates least from the vector x, for the case where G is of finite dimensionality and is generated by the linearly independent vectors

$$g_1, g_2, \ldots, g_n .$$

(The existence of the vector y follows from the theorems in Section 8).

In fact, the vector y is the projection of x upon G, and it follows that this vector must satisfy the equations

$$(1) \qquad (x - y, g_k) = 0 \quad (k = 1, 2, \ldots, n),$$

which, when written out in full, have the form

$$(2)\qquad \begin{cases} \lambda_1 (g_1, g_1) + \lambda_2 (g_2, g_1) + \ldots + \lambda_n (g_n, g_1) = (x, g_1) \\ \lambda_1 (g_1, g_2) + \lambda_2 (g_2, g_2) + \ldots + \lambda_n (g_n, g_2) = (x, g_2) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_1 (g_1, g_n) + \lambda_2 (g_2, g_n) + \ldots + \lambda_n (g_n, g_n) = (x, g_n) \end{cases}$$

and represent a system of linear equations in the unknown coefficients λ_k.

The determinant of the system, i.e.

$$\begin{vmatrix} (g_1, g_1) & (g_2, g_1) \ldots (g_n, g_1) \\ (g_1, g_2) & (g_2, g_2) \ldots (g_n, g_2) \\ \ldots\ldots\ldots\ldots & \\ (g_1, g_n) & (g_2, g_n) \ldots (g_n, g_n) \end{vmatrix} = G (g_1, g_2, \ldots, g_n),$$

is called *Gram's determinant* [6, 7] of the system of vectors $g_1, g_2, \ldots, g_n$.

As H is strictly normalized and the vectors g_i are linearly independent, it follows that, when x is an arbitrary vector, the system (2) has one and only one solution. Whence it follows at once that Gram's determinant of linearly independent vectors can never vanish.

In addition, we wish to determine the expression for the square of the error as to the distance by which the vector y approximates the vector x, i.e. for the quantity

$$\delta^2 = (x - y, x - y).$$

By reason of (1), we have the equation

$$\delta^2 = (x - y, x) - (x - y, y) = (x - y, x) = (x, x) - (y, x)$$

or

$$\lambda_1 (g_1, x) + \lambda_2 (g_2, x) + \ldots + \lambda_n (g_n, x) = (x, x) - \delta^2.$$

If we combine this equation with the system (2) and eliminate $\lambda_1, \lambda_2, \ldots, \lambda_n$, we find that

$$\begin{vmatrix} (g_1, g_1) \ldots (g_n, g_1) & (x, g_1) \\ (g_1, g_2) \ldots (g_n, g_2) & (x, g_2) \\ \ldots\ldots\ldots\ldots\ldots & \\ (g_1, g_n) \ldots (g_n, g_n) & (x, g_n) \\ (g_1, x) \ldots (g_n, x) & (x, x) - \delta^2 \end{vmatrix} = 0 .$$

Whence we obtain

$$\delta^2 = \frac{G (x, g_1, g_2, \ldots, g_n)}{G (g_1, g_2, \ldots, g_n)} .$$

Thus we have found the relation

$$(3)\qquad \min_{a_k} \| x - a_1 g_1 - a_2 g_2 - \ldots - a_n g_n \|^2 = \frac{G (x, g_1, g_2, \ldots, g_n)}{G (g_1, g_2, \ldots, g_n)}$$

From this relation and from

$$G (g_1) = (g_1, g_1) > 0 \qquad (g_1 \neq 0)$$

it follows that Gram's determinant is always $\geqq 0$, where the equality sign holds if and only if linear dependence subsists among the vectors (in particular, if one of the vectors is zero).

This result may be considered as a generalization of Schwarz's inequality, which is equivalent to the inequality

$$G(g_1, g_2) \geqq 0 .$$

Suppose that $m < n$ and that the vectors

$$g_0, g_1, g_2, \ldots, g_n$$

are linearly independent. Then

$$(4_1) \quad \min_{\alpha} \| g_k - \alpha_{k+1} g_{k+1} - \ldots - \alpha_n g_n \| \leqq \min_{\beta} \| g_k - \beta_{k+1} g_{k+1} - \ldots - \beta_m g_m \|$$
$$(k = 0, 1, 2, \ldots, m-1)$$

and

$$(4_2) \quad \min_{\alpha} \| g_m - \alpha_{m+1} g_{m+1} - \ldots - \alpha_n g_n \| \leqq \| g_m \| .$$

Whence, by reason of (3),

$$\frac{G(g_k, g_{k+1}, \ldots, g_n)}{G(g_{k+1}, \ldots, g_n)} \leqq \frac{G(g_k, g_{k+1}, \ldots, g_m)}{G(g_{k+1}, \ldots, g_m)}$$
$$(k = 0, 1, 2, \ldots, m-1)$$

and

$$\frac{G(g_m, g_{m+1}, \ldots, g_n)}{G(g_{m+1}, \ldots, g_n)} \leqq G(g_m)$$

or

$$\frac{G(g_0, g_1, \ldots, g_n)}{G(g_0, g_1, \ldots, g_m)} \leqq \frac{G(g_1, g_2, \ldots, g_n)}{G(g_1, g_2, \ldots, g_m)} \leqq \ldots \leqq \frac{G(g_m, g_{m+1}, \ldots, g_n)}{G(g_m)} \leqq$$
$$\leqq G(g_{m+1}, g_{m+2}, \ldots, g_n) .$$

The inequality thus obtained states that

$$(5) \quad G(g_0, g_1, g_2, \ldots, g_n) \leqq G(g_0, g_1, \ldots, g_m)\, G(g_{m+1}, g_{m+2}, \ldots, g_n) .$$

We will show that the equality sign in (5) (for linearly independent $g_0, \ldots, g_n$) holds if and only if each of the vectors $g_0, \ldots, g_m$ is orthogonal to each of the vectors $g_{m+1}, \ldots, g_n$. In fact, a consideration of the foregoing shows that the equality sign in (5) will occur precisely when all the inequalities (4_1), (4_2) change into the equations

$$\min_{\alpha} \| g_k - \alpha_{k+1} g_{k+1} - \ldots - \alpha_n g_n \| = \min_{\beta} \| g_k - \beta_{k+1} g_{k+1} - \ldots - \beta_m g_m \|$$
$$(k = 0, 1, \ldots, m-1),$$
$$\min_{\alpha} \| g_m - \alpha_{m+1} g_{m+1} - \ldots - \alpha_n g_n \| = \| g_m \| .$$

The last equation is possible only for

$$(g_m, g_i) = 0 \qquad (i = m+1, \ldots, n) ;$$

but when $k = m - 1, m - 2, \ldots, 1, 0$, the foregoing relation gives successively

$$(g_k, g_i) = 0 \qquad (i = m + 1, \ldots, n).$$

This proves the proposition.

From (5), we deduce further that

$$G(g_0, g_1, \ldots, g_n) \leqq (g_0, g_0)(g_1, g_1) \cdots (g_n, g_n). \tag{6}$$

Here the equality sign is valid if and only if

$$(g_i, g_k) = 0 \qquad (i \neq k;\ i, k = 0, 1, \ldots, n).$$

The inequality (6) contains as a special case Hadamard's classical theorem on determinants.

Let

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

where the quantities a_{ik} are, in general, complex numbers. We have

$$|A|^2 = \begin{vmatrix} \sum_1^n |a_{1k}|^2 & \sum_1^n a_{1k}\overline{a_{2k}} & \cdots & \sum_1^n a_{1k}\overline{a_{nk}} \\ \sum_1^n a_{2k}\overline{a_{1k}} & \sum_1^n |a_{2k}|^2 & \cdots & \sum_1^n a_{2k}\overline{a_{nk}} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_1^n a_{nk}\overline{a_{1k}} & \sum_1^n a_{nk}\overline{a_{2k}} & \cdots & \sum_1^n |a_{nk}|^2 \end{vmatrix} = \begin{vmatrix} (g_1, g_1) & (g_1, g_2) & \cdots & (g_1, g_n) \\ (g_2, g_1) & (g_2, g_2) & \cdots & (g_2, g_n) \\ \cdots & \cdots & \cdots & \cdots \\ (g_n, g_1) & (g_n, g_2) & \cdots & (g_n, g_n) \end{vmatrix},$$

where g_k is the vector with the components $a_{k1}, a_{k2}, \ldots, a_{kn}$ in space of n-dimensions.

Because of (6), we have

$$|A|^2 \leqq \sum_1^n |a_{1k}|^2 \sum_1^n |a_{2k}|^2 \cdots \sum_1^n |a_{nk}|^2, \tag{7}$$

where the equality sign holds only in case that

$$\sum_{k=1}^n a_{\lambda k}\overline{a_{\mu k}} = 0 \qquad (\lambda \neq \mu).$$

The relation (7) is known also as Hadamard's inequality [7].

If

$$|a_{ik}| \leqq M$$

then, by virtue of (7), we have

$$|A| \leqq n^{\frac{n}{2}} M^n.$$

14. An Example of Approximation in Hilbert Space. Suppose

$$q,\ p_1,\ p_2, \ldots,\ p_n$$

are numbers in which the real part of each number is greater than $-\frac{1}{2}$, and let no identical numbers be among them. It is required to find

$$\delta^2 = \min_A \int_0^1 | t^q - A_1 t^{p_1} - A_2 t^{p_2} - \ldots - A_n t^{p_n} |^2 \, dt .$$

Here we are dealing with the Hilbert space L^2 (0, 1). By reason of Section 13,

$$\delta^2 = \frac{G(t^q, t^{p_1}, t^{p_2}, \ldots, t^{p_n})}{G(t^{p_1}, t^{p_2}, \ldots, t^{p_n})},$$

and, since

$$(t^\alpha, t^{\bar\beta}) = \int_0^1 t^{\alpha+\bar\beta} dt = \frac{1}{\alpha + \bar\beta + 1},$$

we come to the formulas

$$G(t^{p_1}, t^{p_2}, \ldots, t^{p_n}) = \begin{vmatrix} \frac{1}{p_1+\bar p_1+1} & \cdots & \frac{1}{p_1+\bar p_n+1} \\ \cdots & \cdots & \cdots \\ \frac{1}{p_n+\bar p_1+1} & \cdots & \frac{1}{p_n+\bar p_n+1} \end{vmatrix}$$

and

$$G(t^q, t^{p_1}, t^{p_2}, \ldots, t^{p_n}) = \begin{vmatrix} \frac{1}{q+\bar q+1} & \frac{1}{q+\bar p_1+1} & \cdots & \frac{1}{q+\bar p_n+1} \\ \frac{1}{p_1+\bar q+1} & \frac{1}{p_1+\bar p_1+1} & \cdots & \frac{1}{p_1+\bar p_n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{p_n+\bar q+1} & \frac{1}{p_n+\bar p_1+1} & \cdots & \frac{1}{p_n+\bar p_n+1} \end{vmatrix}.$$

We now evaluate the determinant [8]

$$D_m = \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{a_m+b_1} & \frac{1}{a_m+b_2} & \cdots & \frac{1}{a_m+b_m} \end{vmatrix} = \left| \frac{1}{a_\lambda + b_\mu} \right|_1^m$$

and observe first of all that

$$D_m = \frac{P_m}{\prod_{\lambda,\mu=1}^{m} (a_\lambda + b_\mu)} .$$

Here P_m is a polynomial in $a_1, a_2, \ldots, a_m$ and $b_1, b_2, \ldots, b_m$ of degree $m^2 - m$. On the other hand, P_m will be equal to zero if $a_\lambda = a_\mu$ or $b_\lambda = b_\mu$ for $\lambda \neq \mu$. Therefore P_m is divisible by $A_m B_m$, where

$$A_m = \prod_{1 \leq k < j \leq m} (a_j - a_k), \quad B_m = \prod_{1 \leq k < j \leq m} (b_j - b_k) .$$

But as A_m and B_m are polynomials of degree $\frac{m(m-1)}{2}$ it follows that

$$P_m = a_m A_m B_m ,$$

where a_m is independent of the quantities $a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_m$. It is easy to see that $a_m = 1$. In fact, $a_1 = 1$. On the other hand, by multiplying the last row of D_m by a_m and then allowing first a_m and thereafter b_m to approach ∞, we find that

$$a_m D_m \to D_{m-1}$$

while

$$\lim \frac{a_m \prod_{1 \leq k < j \leq m} (a_j - a_k)(b_j - b_k)}{\prod_{\lambda, \mu = 1}^{m} (a_\lambda + b_\mu)} = \frac{\prod_{1 \leq k < j \leq m-1} (a_j - a_k)(b_j - b_k)}{\prod_{\lambda, \mu = 1}^{m-1} (a_\lambda + b_\mu)}$$

Whence $a_m = a_{m-1}$.

Hence

$$D_m = \frac{\prod_{1 \leq k < j \leq m} (a_j - a_k)(b_j - b_k)}{\prod_{\lambda, \mu = 1}^{m} (a_\lambda + b_\mu)} .$$

If we put

$$a_j = p_j, \quad b_j = \overline{p_j} + 1 ,$$

we find that

$$G(t^{p_1}, t^{p_2}, \ldots, t^{p_n}) = \frac{\prod_{1 \leq k < j \leq n} |p_j - p_k|^2}{\prod_{\lambda, \mu = 1}^{n} (p_\lambda + \overline{p_\mu} + 1)}$$

and analogously

$$G(t^q, t^{p_1}, t^{p_2}, \ldots, t^{p_n}) = \frac{\prod_{k=1}^{n} |q - p_k|^2 \prod_{1 \leq k < j \leq n} |p_j - p_k|^2}{(q + \overline{q} + 1) \prod_{k=1}^{n} |q + \overline{p_k} + 1|^2 \prod_{\lambda, \mu = 1}^{n} (p_\lambda + \overline{p_\mu} + 1)}$$

Therefore

$$\delta^2 = \frac{1}{q + \overline{q} + 1} \prod_{k=1}^{n} \left| \frac{\overline{q} - p_k}{q + p_k + 1} \right|^2 .$$

15. More About the Approximation Problem in Hilbert Space. We will consider the complete Hilbert space H and assume as given therein a closed subspace G [11] (which need not have a finite number of dimensions; the latter is always closed). We will prove that, for any vector $x \in H$, there exists one and (because the space is strictly normalized) only one vector $y \in G$ which, among all the vectors of the subspace G, furnishes the best approximation of x.

For that purpose, let us denote the quantity

$$d = \inf_{g \in G} || x - g ||$$

by d, and prove first that, for any two vectors $g', g'' \in G$, the inequality

$$|| g' - g'' || \leqq \sqrt{|| x - g' ||^2 - d^2} + \sqrt{|| x - g'' ||^2 - d^2}$$

holds. This inequality is called the generalized inequality of Beppo Levi. In the case of ordinary three-dimensional space, the said inequality follows from the Pythagorean theorem (Fig. 2).

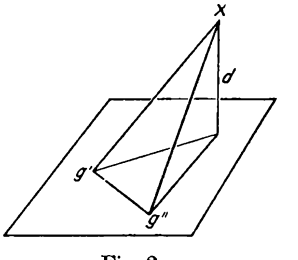


Fig. 2

For the proof of Beppo Levi's inequality, we make use of the fact that, for any λ $(\neq -1)$,

$$g = \frac{g' + \lambda g''}{1 + \lambda} \in G$$

where λ is assumed real. Therefore

$$|| x - g ||^2 \geqq d^2$$

and hence

$$|| (x - g') + \lambda (x - g'') ||^2 \geqq d^2 (1 + \lambda)^2$$

or

$$\lambda^2 \{|| x - g'' ||^2 - d^2\} + 2\lambda \{\Re (x - g'', x - g') - d^2\} + \\ + \{|| x - g' ||^2 - d^2\} \geqq 0.$$

Since, by hypothesis, λ is arbitrary and real, we must have

$$\{\Re (x - g'', x - g') - d^2\}^2 \leqq \{|| x - g' ||^2 - d^2\} \{|| x - g'' ||^2 - d^2\} .$$

Consequently,

$$\begin{aligned} || g' - g'' ||^2 &= || (x - g') - (x - g'') ||^2 = \\ &= || x - g' ||^2 - 2\Re (x - g', x - g'') + || x - g'' ||^2 = \\ &= \{|| x - g' ||^2 - d^2\} - 2 \{\Re (x - g', x - g'') - d^2\} + \\ &\quad + \{|| x - g'' ||^2 - d^2\} \leqq \\ &\leqq \{|| x - g' ||^2 - d^2\} + 2 \sqrt{\{|| x - g' ||^2 - d^2\} \{|| x - g'' ||^2 - d^2\}} + \\ &\quad + \{|| x - g'' ||^2 - d^2\} = \\ &= \left\{\sqrt{|| x - g' ||^2 - d^2} + \sqrt{|| x - g'' ||^2 - d^2}\right\}^2. \end{aligned}$$

Thus the inequality of Beppo Levi has been proved.

[11] Translator's note: It is not assumed here that generally subspaces of the Hilbert space H are closed.

We now select from G a sequence

$$g_1, g_2, g_3, \ldots,$$

such that

$$\lim_{n \to \infty} || x - g_n || = d \cdot$$

Then, by virtue of Beppo Levi's inequality,

$$\lim_{m, n \to \infty} || g_n - g_m || = 0,$$

that is to say, the sequence

$$g_1, g_2, g_3, \ldots$$

converges by Cauchy's test, and, because H is a complete space, there exists a vector $y \in H$ to which the given sequence converges. This vector belongs to G, since by hypothesis G is closed.

We are going to prove that

$$|| x - y || = d \cdot$$

In fact, we have

$$|| x - y || \leqq || x - g_n || + || y - g_n ||$$

and consequently also

$$|| x - y || \leqq d.$$

But since $y \in G$, we have

$$|| x - y || \geqq d.$$

Whence it follows that

$$|| x - y || = d$$

and our theorem is proved [9].

16. Orthonormalized Vector Systems in Hilbert Space. We are going to show that the law of formation of the linear expression

$$\lambda_1 g_1 + \lambda_2 g_2 + \ldots + \lambda_n g_n ,$$

when it represents the shortest distance to x, assumes a particularly simple form if the vectors g_i satisfy the relations

$$(g_i, g_k) = \begin{cases} 0 & (i \neq k) \\ 1 & (i = k) \end{cases} \qquad (i, k = 1, 2, \ldots, n) \cdot$$

We say of such vectors, that they constitute an *orthonormalized* system. In fact, the system of equations for the unknown coeffients λ_i in Section 13 now assumes the form

$$\lambda_k (g_k, g_k) = (x, g_k) \qquad (k = 1, 2, \ldots, n),$$

and we obtain

$$\lambda_k = (x, g_k) \qquad (k = 1, 2, \ldots, n) ;$$

that is to say, every coefficient λ_k can be calculated independently of the rest by means of a very simple rule.

If an orthonormalized system of vectors

$$g_1, g_2, \ldots, g_n$$

can be extended to the orthonormalized system

$$g_1, g_2, \ldots, g_n, g_{n+1}$$

we see that the new expression for the best approximation is found by adding an extra term

$$(x, g_{n+1})\, g_{n+1} ,$$

there being no change in the coefficients of the old expression.

In the case where the system

$$g_1, g_2, \ldots, g_n$$

is orthonormalized, the square of the error in the best approximation of every vector x assumes a very simple form. In fact,

$$\delta^2 = \left\| x - \sum_{k=1}^{n} (x, g_k)\, g_k \right\|^2 = \tag{1}$$

$$= \| x \|^2 - \left(x, \sum_{k=1}^{n} (x, g_k)\, g_k \right) - \left(\sum_{k=1}^{n} (x, g_k)\, g_k, x \right) + \left\| \sum_{k=1}^{n} (x, g_k)\, g_k \right\|^2 =$$

$$= \| x \|^2 - \sum_{k=1}^{n} | (x, g_k) |^2 .$$

17. Orthogonalization of Vector Systems. Two vector systems

$$g_1, g_2, \ldots, g_n$$
$$g_1', g_2', \ldots, g_n'$$

are said to be *equivalent* if they span one and the same subspace G. This means that every vector of the one system can be represented as a linear combination of the vectors of the other system. It is easy to see that the necessary and sufficient condition for this is

$$g_k' = a_{1k} g_1 + a_{2k} g_2 + \ldots + a_{nk} g_n \qquad (k = 1, 2, \ldots, n)$$

with

$$\begin{vmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots & \ldots \ldots \\ a_{n1} & a_{n2} \ldots a_{nn} \end{vmatrix} \neq 0 .$$

It is clear that in the problems of approximation in which we are interested, a system can always be replaced by an equivalent system.

The previous section leads naturally to the following query: Given, in H, the linear independent vectors

$$g_1, g_2, \ldots, g_n ; \tag{1}$$

is it possible to find an equivalent orthogonal vector system

$$e_1, e_2, \ldots, e_n ?$$

In other words, can the system (1) be orthogonalized? It turns out that the orthogonalization is always possible.

This can be done by a simple method due to E. Schmidt [10], which we will now explain. Let us put

$$e_1 = \frac{g_1}{\| g_1 \|},$$

and recall that $\| g_1 \| \neq 0$ by virtue of the linear independence of the vectors (1). Clearly $\| e_1 \| = 1$. Consider also the vector

$$g_2' = g_2 - \lambda_1 e_1 .$$

Here λ_1 can be so chosen that the vector g_2' is orthogonal to e_1. Such is the case for

$$\lambda_1 = (g_2, e_1) .$$

Consequently

$$g_2' = g_2 - (g_2, e_1)\, e_1 ,$$

where $\| g_2' \| \neq 0$, since otherwise the vectors g_2 and e_1 and hence also the vectors g_2 and g_1 would be linearly dependent, contrary to hypothesis.

If we put

$$e_2 = \frac{g_2'}{\| g_2' \|},$$

then

$$\| e_2 \| = 1, \quad (e_2, e_1) = 0 .$$

Next we construct the vector

$$g_3' = g_3 - (g_3, e_1)\, e_1 - (g_3, e_2)\, e_2 ,$$

which is orthogonal to the vectors e_1, e_2 and obviously different from zero, and put

$$e_3 = \frac{g_3'}{\| g_3' \|} .$$

It is clear that

$$\| e_3 \| = 1, \quad (e_3, e_1) = 0, \quad (e_3, e_2) = 0 .$$

Let us put generally

$$g_k' = g_k - (g_k, e_1)\, e_1 - (g_k, e_2)\, e_2 - \ldots - (g_k, e_{k-1})\, e_{k-1},$$

$$e_k = \frac{g_k'}{\| g_k' \|}.$$

The system

$$e_1, e_2, \ldots, e_n,$$

is evidently orthogonal and normalized. It is equivalent to the system (1) because the following equations are satisfied:

$$\begin{aligned} e_1 &= a_{11} g_1, \\ e_2 &= a_{21} g_1 + a_{22} g_2, \\ &\cdots\cdots\cdots\cdots \\ e_n &= a_{n1} g_1 + a_{n2} g_2 + \ldots + a_{nn} g_n \end{aligned}$$

and

$$a_{kk} \neq 0 \qquad (k = 1, 2, \ldots, n).$$

The reader can easily verify that the explicit expression for the vectors e_i is given by the formulas

$$e_k = \frac{1}{\sqrt{G_k G_{k-1}}} \begin{vmatrix} (g_1, g_1) & (g_2, g_1) & \ldots & (g_k, g_1) \\ (g_1, g_2) & (g_2, g_2) & \ldots & (g_k, g_2) \\ \cdots & \cdots & \cdots & \cdots \\ (g_1, g_{k-1}) & (g_2, g_{k-1}) & \ldots & (g_k, g_{k-1}) \\ g_1 & g_2 & \ldots & g_k \end{vmatrix},$$

$$G_0 = 1, \quad G_k = G(g_1, g_2, \ldots, g_k)$$
$$(k = 1, 2, \ldots, n).$$

18. Infinite Orthonormalized Systems. In three-dimensional Euclidean space, a vector which is perpendicular to all three vectors of an orthonormalized system is of course equal to the zero vector. In other words: In this case it is impossible to extend an orthonormalized system consisting of three vectors; it is manifestly complete.

An orthonormalized system in H is said to be *complete* if there does not exist any vector, other than the zero vector, orthogonal to every vector of the system in question.

We will show that in a separable Hilbert space the vectors of all orthonormalized systems can be numbered, that is, they form either finite or denumerable sets.

Let $\{x_k\}$ be a denumerable set, dense in H. Let e' and e'' be two distinct given vectors of our orthonormalized system. Let us associate with each of these vectors an element of $\{x_k\}$ which is distant less than $\frac{1}{2}\sqrt{2}$ from its corresponding vector. I.e., let

$$\| x_{k'} - e' \| < \frac{1}{2}\sqrt{2}, \quad \| x_{k''} - e'' \| < \frac{1}{2}\sqrt{2}.$$

If we show that $k' \neq k''$, then no more than a single vector of our orthonormalized system can correspond to each element of $\{x_k\}$, and our proposition will be proved. Suppose $k' = k'' = k$. Then, on the one hand, the triangular inequality gives

$$\| e' - e'' \| \leqq \| e' - x_k \| + \| e'' - x_k \| < \sqrt{2},$$

and on the other hand

$$\| e' - e'' \|^2 = (e' - e'', e' - e'') = \| e' \|^2 + \| e'' \|^2 = 2,$$

which is absurd. This proves the proposition.

If an infinite system of orthonormalized vectors

$$e_1, e_2, e_3, \ldots$$

in the (separable or non-separable) space H is given, we can associate with every vector $x \in H$ a series

$$\sum_{k=1}^{\infty} (x, e_k)\, e_k, \tag{1}$$

that is, a sequence of vectors

$$\sum_{k=1}^{n} (x, e_k)\, e_k.$$

Since, by virtue of (1), Section 16,

$$\sum_{k=1}^{n} | (x, e_k) |^2 \leqq (x, x),$$

the series

$$\sum_{k=1}^{\infty} | (x, e_k) |^2$$

converges, and we have the inequality

$$\sum_{k=1}^{\infty} | (x, e_k) |^2 \leqq (x, x). \tag{2}$$

In analogy to trigonometric series, we call the series (1) *Fourier's series* for x and express the correspondence as

$$x \sim \sum_{k=1}^{\infty} (x, e_k)\, e_k.$$

Formula (2) is called *Bessel's inequality.*

From the expression in section 16 for the square of the deviation in the best approximation, it follows that the vector $x \in H$ can be approximated with an arbitrary accuracy by means of expressions of the form

$$a_1 e_1 + a_2 e_2 + \ldots + a_n e_n$$

if and only if Bessel's inequality for the vector x is changed into the equation

$$\sum_{k=1}^{\infty} |(x, e_k)|^2 = (x, x),$$

which we designate as *Parseval's equation* (again in analogy to trigonometric series).

THEOREM: *In order that the orthonormalized sequence*

$$e_1, e_2, e_3, \ldots \tag{3}$$

in H be complete, it is sufficient, and in the case where the space H is complete [12] *it is also necessary, that Parseval's equation*

$$(x, x) = \sum_{k=1}^{\infty} |(x, e_k)|^2$$

shall be valid for all vectors $x \in H$.

Proof: Suppose the sequence (3) in H is not complete. Then there exists a vector $x \in H$ such that

$$\|x\| = 1, \quad (x, e_k) = 0 \qquad (k = 1, 2, 3, \ldots).$$

Hence

$$1 = (x, x) > \sum_{k=1}^{\infty} |(x, e_k)|^2 = 0,$$

and consequently Parseval's equation is not true for all the vectors. Thus we have shown that the condition of the theorem is sufficient.

Let us now assume that the space is complete and that (3) is a complete orthonormalized sequence therein. We must then show that Parseval's equation is true for all vectors $x \in H$. In the contrary case, there would exist a vector x such that

$$(x, x) > \sum_{k=1}^{\infty} |(x, e_k)|^2 .$$

Let us now consider the sequence of vectors

$$x_n = \sum_{k=1}^{n} (x, e_k)\, e_k \qquad (k = 1, 2, 3, \ldots). \tag{4}$$

[12] The concepts "completeness of a space" and "complete orthonormalized systems" are to be distinguished. — Translators note.

Since we have

$$|| x_n - x_m ||^2 = \sum_{k=m+1}^{n} | (x, e_k) |^2 \to 0 .$$

when $m, n \to \infty (n > m)$, it follows that the sequence $\{x_n\}$ converges in the sense of Cauchy and hence there exists a vector $x \in H$ (by hypothesis, the space H is complete), such that

$$|| x' - x_n || \to 0 . \tag{5}$$

By virtue of Schwarz's inequality we get

$$\lim_{n \to \infty} (x' - x_n, e_k) = 0.$$

But when $n \leq k$, we have

$$(x' - x_n, e_k) = (x', e_k) - (x, e_k).$$

Whence

$$(x', e_k) = (x, e_k) \qquad (k = 1, 2, 3, \ldots). \tag{6}$$

Take the vector $y = x - x'$. By reason of (6), it is orthogonal to all the vectors of the sequence (3). Consequently, it must be zero, because, by hypothesis, the system (3) is complete. That is impossible for, on the one hand,

$$|| x ||^2 > \sum_{k=1}^{\infty} | (x, e_k) |^2$$

and, on the other hand, by reason of (4), (5), (6), the equality

$$|| x' ||^2 = \sum_{k=1}^{\infty} | (x, e_k) |^2$$

holds, at the same time; and hence we get the inequality

$$|| y || \geqq || x || - || x' || > 0.$$

In conclusion, we will show that the existence of Parseval's equation for all vectors $x \in H$ implies the generalized Parseval equation

$$(x, y) = \sum_{k=1}^{\infty} (x, e_k) (e_k, y) \tag{7}$$

for every pair of vectors $x, y \in H$.

Applying Parseval's equation to the vector $x + \lambda y$, we get

$$(x, x) + \lambda (y, x) + \bar{\lambda} (x, y) + |\lambda|^2 (y, y) = \sum_{k=1}^{\infty} |(x + \lambda y, e_k)|^2 =$$

$$= \sum_{k=1}^{\infty} |(x, e_k) + \lambda (y, e_k)|^2 =$$

$$= \sum_{k=1}^{\infty} |(x, e_k)|^2 + \lambda \sum_{k=1}^{\infty} (y, e_k)(e_k, x) + \bar{\lambda} \sum_{k=1}^{\infty} (x, e_k)(e_k, y) +$$

$$+ |\lambda|^2 \sum_{k=1}^{\infty} |(y, e_k)|^2 .$$

Whence

$$\lambda (y, x) + \bar{\lambda} (x, y) = \lambda \sum_{k=1}^{\infty} (y, e_k)(e_k, x) + \lambda \sum_{k=1}^{\infty} (x, e_k)(e_k, y) .$$

If we substitute for λ first 1 and then i, we obtain the relation (7), which was to be proved.

19. An Example of a Non-separable Space. The totality of Bohr's almost-periodic functions $x(t)$, $y(t)$, ... constitutes a certain Hilbert space, if the inner product is defined by means of the formula

$$(x, y) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \bar{y}(t) \, dt .$$

The substance of one of the fundamental theorems of Bohr's theory [11] is that the equation $x(t) = 0$ follows from $(x, x) = 0$.

The space under consideration is not separable. This follows from the fact that there exists in it a set of the power of the continuum of linearly independent orthonormalized vectors. As a matter of fact, the functions $e^{i\alpha t}$ possess this property (α is any real number). Each of these functions is purely periodic and hence belongs to the space; on the other hand,

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} e^{i\alpha t} e^{-i\beta t} \, dt = \begin{cases} 0 & (\alpha \neq \beta), \\ 1 & (\alpha = \beta). \end{cases}$$

20. Weierstrass' First Theorem [12]. From the theorem of Section 8 and its application to the space C, we obtain the following. Let $f(x)$ be a function which is continuous in the finite interval $[a, b]$; then, for any value of n, there exists a polynomial

$$P_n(x) = p_0 x^n + p_1 x^{n-1} + \ldots + p_n ,$$

whose deviation from $f(x)$ is less than that of any other polynomial, in the sense that

$$\varrho_n = \max_{a \leq x \leq b} |f(x) - P_n(x)| \leqq \max_{a \leq x \leq b} |f(x) - Q_n(x)| ,$$

where $Q_n(x)$ is any polynomial of the n-th degree.

The inequality $\varrho_{n+1} \leqq \varrho_n$ follows immediately.

We will now prove that $\varrho_n \to 0$ when $n \to \infty$. This assertion is the substance of Weierstrass' theorem (1885), which says: *If $f(x)$ is a function which is continuous in the finite interval $[a, b]$, then for every $\varepsilon > 0$ there exists a polynomial $P_n(x)$ of degree $n = n(\varepsilon)$ such that the inequality*

$$|f(x) - P_n(x)| \leqq \varepsilon$$

holds throughout the interval $[a, b]$.

Without loss of generality, we may take $a = 0$, $b = 1$. The proof we are about to give is due to S. N. Bernstein [13].

This proof is accomplished by constructing the polynomial

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$$

($n \geqq 1$) and then showing that

$$\lim_{n \to \infty} B_n(x) = f(x)$$

uniformly throughout the interval $[0, 1]$.

First of all, we have the following identities:

(1)
$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1,$$

$$\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} = x,$$

$$\sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} x^k (1-x)^{n-k} = \left(1 - \frac{1}{n}\right) x^2 + \frac{1}{n} x.$$

The last two identities are found by differentiating the relation

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n$$

with respect to p. From the above identities, it follows that

(2)
$$\sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1-x)^{n-k} = \frac{x(1-x)}{n}.$$

Multiplying (1) by $f(x)$ and subtracting $B_n(x)$, we get

$$f(x) - B_n(x) = \sum_{k=0}^{n} \left\{ f(x) - f\left(\frac{k}{n}\right) \right\} \binom{n}{k} x^k (1-x)^{n-k} = \sum' + \sum''.$$

In the last expression, the summation in $\sum'$ runs through the values of k for which

$$\left|\frac{k}{n} - x\right| \leqq \frac{1}{\sqrt[4]{n}}$$

and the summation in $\sum''$ runs through the remaining values of k.

As $f(x)$ is continuous in the interval $[0, 1]$ and is therefore bounded, that is, $|f(x)| \leq M$ throughout the interval, it follows that

$$\left|\sum''\right| \leqq 2M \sum'' \binom{n}{k} x^k (1-x)^{n-k} =$$

$$= 2M \sum'' \frac{(k-nx)^2}{(k-nx)^2} \binom{n}{k} x^k (1-x)^{n-k} \leqq$$

$$\leqq 2M \sum'' \frac{(k-nx)^2}{\sqrt{n^3}} \binom{n}{k} x^k (1-x)^{n-k} \leqq$$

$$\leqq \frac{2M}{\sqrt{n^3}} \sum_{k=0}^{n} (k-nx)^2 \binom{n}{k} x^k (1-x)^{n-k};$$

from this relation and by virtue of (2), we get

$$\left|\sum''\right| \leqq \frac{2M}{\sqrt{n^3}} nx(1-x) \leqq \frac{M}{2\sqrt{n}}.$$

On the other hand, we have the inequalities

$$\left|\sum'\right| \leqq \varepsilon_n \sum' \binom{n}{k} x^k (1-x)^{n-k} \leqq \varepsilon_n \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \varepsilon_n,$$

where

$$\varepsilon_n = \max_{\left|\frac{k}{n} - x\right| \leq \frac{1}{\sqrt[4]{n}}} \left| f(x) - f\left(\frac{k}{n}\right) \right|.$$

Consequently, $\varepsilon_n \to 0$ as $n \to \infty$.[13]

Finally,

$$|f(x) - B_n(x)| \leqq \varepsilon_n + \frac{M}{2\sqrt{n}},$$

and thus the proof of Weierstrass' theorem is complete.

Remark: If $P_n(x)$ approaches $f(x)$ uniformly as $n \to \infty$, then $f(x)$ can be developed in the uniformly convergent series

$$f(x) = P_1(x) + \{P_2(x) - P_1(x)\} + \{P_3(x) - P_2(x)\} + \ldots$$

---

[13] Uniformly for every value of x in the interval $[0, 1]$. — Translator's note.

Therefore Weierstrass' theorem declares further, that every function $f(x)$ which is continuous in a finite closed interval $[a, b]$ can be developed according to polynomials in a series which is uniformly convergent for $a \leqq x \leqq b$.

21. Weierstrass' Second Theorem [12]. This theorem refers to periodic continuous functions and says: *If $F(t)$ is a continuous function of period 2π, then, corresponding to every positive number ε there exists a trigonometric sum*

$$S_n(t) = a_0 + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt)$$

$[n = n(\varepsilon)]$ *such that the inequality*

$$|F(t) - S_n(t)| \leqq \varepsilon$$

is satisfied for all values of t.

To prove the theorem, let us consider the two functions

$$\varphi(t) = \frac{F(t) + F(-t)}{2}, \quad \psi(t) = \frac{F(t) - F(-t)}{2} \sin t.$$

They are continuous even functions of period 2π. Let us set $x = \cos t$, and consider the interval $[0, \pi]$ for the variable t, or, what comes to the same thing, the interval $[-1, 1]$ for the variable x. We see that

$$\varphi(t) = f(x), \qquad \psi(t) = g(x)$$

are functions of x which are continuous in the interval $[-1, 1]$. Therefore, by reason of Weierstrass' first theorem, there exist polynomials $P(x)$, $Q(x)$ which satisfy the relations

$$|f(x) - P(x)| \leqq \frac{\varepsilon}{4}, \qquad |g(x) - Q(x)| \leqq \frac{\varepsilon}{4} \qquad (-1 \leqq x \leqq 1),$$

or

$$|\varphi(t) - P(\cos t)| \leqq \frac{\varepsilon}{4}, \qquad |\psi(t) - Q(\cos t)| \leqq \frac{\varepsilon}{4}.$$

These inequalities were obtained in the first instance for the interval $0 \leqq t \leqq \pi$. But they are valid for all values of t, since all the functions appearing therein are periodic and even.

Observing that

$$F(t) \sin t = \varphi(t) \sin t + \psi(t)$$

we obtain for all values of t

$$|F(t) \sin t - U(t)| \leqq \frac{\varepsilon}{2}, \tag{1}$$

where

$$U(t) = Q(\cos t) + P(\cos t) \sin t$$

is a certain trigonometric sum.

If we apply the same reasoning to the function $F\left(\frac{\pi}{2} - t\right)$, we obtain in a similar fashion

$$\left| F\left(\frac{\pi}{2} - t\right) \sin t - V(t) \right| \leqq \frac{\varepsilon}{2} ,$$

where $V(t)$ is a certain other trigonometric sum.

Substituting t for $\frac{\pi}{2} - t$ in the last inequality, it becomes

$$\left| F(t) \cos t - V\left(\frac{\pi}{2} - t\right) \right| \leqq \frac{\varepsilon}{2} . \tag{2}$$

It follows from (1) and (2), that

$$| F(t) \sin^2 t - U(t) \sin t | \leqq \frac{\varepsilon}{2} ,$$

$$\left| F(t) \cos^2 t - V\left(\frac{\pi}{2} - t\right) \cos t \right| \leqq \frac{\varepsilon}{2} .$$

Therefore

$$\left| F(t) - U(t) \sin t - V\left(\frac{\pi}{2} - t\right) \cos t \right| \leqq \varepsilon ,$$

and consequently the trigonometric sum

$$S(t) = U(t) \sin t + V\left(\frac{\pi}{2} - t\right) \cos t$$

fulfills all our requirements.

22. The Separability of the Space C. We are going to show that the space C (with respect to a finite interval) is separable.

According to Weierstrass' theorem, the aggregate of all polynomials is dense in C. If we now replace the (in general, complex) coefficients of any one of the polynomials by others sufficiently near these given numbers and having rational real and imaginary parts (these will be called briefly rational complex numbers), we shall obtain a polynomial which differs from the original by as little as we please.

Thus the aggregate of all polynomials with rational complex coefficients is also dense in C. But it is easy to see that the set of all polynomials with rational coefficients is denumerable. This follows from the fact that we can pair each of these polynomials with a number

$$n + q + \sum_{k=0}^{n} (| p_k' | + | p_k'' |)$$

called the "weight." Here, n is the degree of the polynomial, the positive integer q is the common denominator of the coefficients, and $p_k' + ip_k''$ is the numerator

of the coefficients of the k-th power. After the coefficients have been reduced to a common denominator, there can be only a finite number of polynomials of given weight.

Since a closed infinite interval can be mapped in a one-to-one manner on a finite interval, it follows that the space C is also separable with respect to a closed infinite interval.

23. The Separability of the Space L^p. In the Lebesgue theory of integrals, the following statements are proven:

a) If the function $x(t)$ is integrable in the interval $[a, b]$ (the interval is assumed to be finite) and if we put

$$z(t) = \int_a^t x(\tau)\, d\tau \qquad (a \leqq t \leqq b),$$

then we have

$$z'(t) = x(t).$$

almost everywhere in $[a, b]$.

b) If a sequence of measurable functions $\{x_n(t)\}$ is uniformly bounded and if, almost everywhere,

$$\lim_{n \to \infty} x_n(t) = x(t),$$

then we have

$$\lim_{n \to \infty} \int_a^b x_n(t)\, dt = \int_a^b x(t)\, dt.$$

c) If $x(t) \in L^p$ when $p > 1$, then $x(t) \in L$.

The first two of these statements will be assumed without proof [14].

The last statement can be easily proved as follows: Put

$$x(t) = x_1(t) + x_2(t),$$

where

$$x_1(t) = \begin{cases} x(t), & \text{for } |x(t)| > 1, \\ 0, & \text{for } |x(t)| \leqq 1. \end{cases}$$

Then the function $x_2(t)$ is integrable, since $|x_2(t)| \leq 1$, and the integrability of $x_1(t)$ follows from

$$|x_1(t)| \leqq |x(t)|^p \in L.$$

Now we come to the proof of the separability of the space L^p $(p \geq 1)$. Let the interval of integration be $[0, 1]$; we will denote it by I. Let us put

$$u_r^{(n)}(t) = \begin{cases} 1 \text{ for } \dfrac{r}{n} \leqq t < \dfrac{r+1}{n}, \\ 0 \text{ otherwise} \end{cases}$$

$$(r = 0, 1, \ldots, n-1; \qquad n = 1, 2, 3, \ldots).$$

The set of functions

$$\sum_{r=0}^{n-1} \lambda_r^{(n)} u_r^{(n)}(t) \in L^p \qquad (n = 1, 2, \ldots) \tag{1}$$

with the rational coefficients $\lambda_r^{(n)}$, is obviously denumerable. We are going to prove that this set is dense in L^p. To that end, it suffices to show that the set of functions (1) with the arbitrary complex coefficients $\lambda_r^{(n)}$ is dense in L^p.

Suppose $x(t) \in L^p$. Let us put

$$s_n(t) = n \sum_{r=0}^{n-1} u_r^{(n)}(t) \int_{\frac{r}{n}}^{\frac{r+1}{n}} x(\tau)\, d\tau,$$

wherein, by virtue of proposition c), the integral on the right-hand side exists. For any $t \in I$, we evidently have

$$s_n(t) = n \int_{\frac{[nt]}{n}}^{\frac{[nt]+1}{n}} x(\tau)\, d\tau = \frac{z\left(\frac{[nt]+1}{n}\right) - z\left(\frac{[nt]}{n}\right)}{\frac{1}{n}}$$

with

$$z(t) = \int_0^t x(\tau)\, d\tau.$$

On the basis of statement a), we have

$$\lim_{h,\, k \to 0} \frac{z(t+h) - z(t-k)}{h+k} = x(t)$$

almost everywhere in I. Whence we have, almost everywhere in I,

$$\lim_{n \to \infty} s_n(t) = x(t).$$

Let us assume first that $x(t)$ is a bounded function ($|x(t)| < M$). Then we have

$$\left| \int_{\frac{[nt]}{n}}^{\frac{[nt]+1}{n}} x(t)\, dt \right| < \frac{M}{n}, \quad |s_n(t)| < M$$

and consequently

$$|x(t) - s_n(t)| < 2M.$$

Therefore, by virtue of statement b),

$$\lim_{n \to \infty} \int_0^1 | x(t) - s_n(t) |^p \, dt = 0,$$

that is to say,

$$\int_0^1 | x(t) - s_n(t) |^p \, dt < \varepsilon$$

for sufficiently large values of n.

Let us now assume that $x(t)$ is not bounded. Suppose that E_k is the set of those points, in the interval I, at which $| x(t) | > k$ $(k = 1, 2, 3, \ldots)$. Let us put

$$x_k(t) = \begin{cases} x(t) & (t \in I - E_k), \\ 0 & (t \in E_k); \end{cases}$$

then we shall have

$$\int_0^1 | x(t) - x_k(t) |^p \, dt = \int_{E_k} | x(t) |^p \, dt,$$

and the right-hand side approaches 0 as $k \to \infty$. This fact reduces the study of unbounded functions $x(t)$ to the study of functions which are bounded for sufficiently large values of k.

Thus we have proved that the space $L^p(0, 1)$ is separable.

It is easy to pass from the space $L^p(0, 1)$ to the space $L^p(a, b)$ for any finite or infinite values of a, b. Since $L^p(-\infty, \infty)$ contains all the spaces $L^p(a, b)$ as subspaces (if we define all the functions $f(t) \in L^p(a, b)$ along the entire number axis and equate them to zero outside the interval (a, b)), it suffices to consider $L^p(-\infty, \infty)$. Let

$$t = \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} u \qquad (-\infty < u < \infty)$$

and

$$g(u) = \frac{1}{\sqrt[p]{\pi}} \frac{f(t)}{(1 + u^2)^{\frac{1}{p}}}.$$

Now, $f(t) \in L^p(0, 1)$ implies $g(u) \in L^p(-\infty, \infty)$, and conversely; we have therewith

$$\int_{-\infty}^{\infty} | g(u) |^p \, du = \int_0^1 | f(t) |^p \, dt.$$

Whence it follows that the sequences

$$\{ f_k(t) \}_{k=1}^{\infty}$$

fill $L^p(0, 1)$ densely; the same thing is true of the sequences

$$\{ g_k(u) \}_{k=1}^{\infty}$$

with respect to $L^p(-\infty, \infty)$.

24. Generalization of Weierstrass' Theorem to the Space L^p.

THEOREM: *If* $x(t) \in L^p(0, 1)$ [14] $(p \geq 1)$, *then, corresponding to every number* $\varepsilon > 0$, *there exists a polynomial* $P(t)$ *such that*

$$\| x - P \|_p = \left\{ \int_0^1 | x(t) - P(t) |^p \, dt \right\}^{\frac{1}{p}} \leqq \varepsilon \ .$$

By reason of the previous section, it will suffice to consider the case

$$x(t) = \begin{cases} 1 & \alpha \leqq t \leqq \beta \quad (0 \leqq \alpha < \beta \leqq 1), \\ 0 & \text{in } I = [0, 1] \ . \end{cases}$$

Let us construct a continuous function $y(t)$ which is: equal to 1 when $\alpha \leqq t \leqq \beta$; equal to 0 when $0 \leqq t \leqq \alpha - \left(\frac{1}{4}\varepsilon\right)^p$ and $\beta + \left(\frac{1}{4}\varepsilon\right)^p \leqq t \leqq 1$; linear in the rest of I. (It is assumed here that $\alpha > \left(\frac{1}{4}\varepsilon\right)^p$, $(1 - \beta) > \left(\frac{1}{4}\varepsilon\right)^p$. If these conditions are not fulfilled, it is obviously necessary to change slightly the construction of the function $y(t)$.)

We have

$$\int_0^1 | x(t) - y(t) |^p \, dt \leqq 2\left(\frac{1}{4}\varepsilon\right) \leqq \left(\frac{1}{2}\varepsilon\right)^p,$$

and since $y(t)$ is continuous, there exists by Weierstrass' theorem a polynomial $P(t)$ such that

$$| y(t) - P(t) | \leqq \frac{1}{2}\varepsilon \qquad (0 \leqq t \leqq 1) \, .$$

Hence, by virtue of Minkowski's inequality, we find

$$\left\{ \int_0^1 | x(t) - P(t) |^p \, dt \right\}^{\frac{1}{p}} \leqq$$

$$\leqq \left\{ \int_0^1 | x(t) - y(t) |^p \, dt \right\}^{\frac{1}{p}} + \left\{ \int_0^1 | y(t) - P(t) |^p \, dt \right\}^{\frac{1}{p}} \leqq \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

[14] Instead of $L^p(0, 1)$ we can obviously take $L^p(a, b)$ for any finite values of a, b.

Evidently, the following proposition is also true: *If* $x(t) \in L^p$ $(p \geq 1)$ *with respect to the interval* $[0, 2\pi]$, *then there exists for every* $\varepsilon > 0$ *a trigonometric sum* $S(t)$ *such that*

$$|| x - S ||_p \leqq \varepsilon .$$

25. The Completeness of the Space L^p. In proving the completeness of the space L^p $(p \geq 1)$ we shall follow J. v. Neumann [15].

Suppose that the sequence of functions $x_n(t) \in L^p$ $(n = 1, 2, \ldots)$ is convergent in the sense of Cauchy:

$$\lim_{m, n \to \infty} \int_{-\infty}^{\infty} | x_n(t) - x_m(t) |^p \, dt = 0.$$

This means that, corresponding to every $\varepsilon > 0$, there exists an $N = N(\varepsilon)$ such that the inequality

$$\int_{-\infty}^{\infty} | x_n(t) - x_m(t) |^p \, dt < \varepsilon$$

holds when $m, n > N$.

Let us construct an infinite sequence of increasing natural numbers N_r $(r = 1, 2, \ldots)$ so that

$$N_r > N\left(\frac{1}{2^{rp+r}}\right).$$

Then we shall have

$$\int_{-\infty}^{\infty} | x_{N_{r+1}} - x_{N_r} |^p \, dt < \frac{1}{2^{rp+r}}.$$

Therefore the set of those points of the interval $I = [-\infty, \infty]$ for which

$$| x_{N_{r+1}}(t) - x_{N_r}(t) | \leq \frac{1}{2^r}$$

has a measure which is less than $\frac{1}{2^r}$. Hence the inequalities

$$| x_{N_{r+1}}(t) - x_{N_r}(t) | < \frac{1}{2^r},$$

$$| x_{N_{r+2}}(t) - x_{N_{r+1}}(t) | < \frac{1}{2^{r+1}}$$

hold simultaneously for the point set I_r. Its complement, with respect to the interval I, has the measure

$$m(I - I_r) < \sum_{k=r}^{\infty} \frac{1}{2^k} = \frac{1}{2^{r-1}}.$$

It will be observed that $I_r \subset I_{r+1} \subset I_{r+2} \ldots$. Hence $\lim_{r\to\infty} I_r = I^*$ exists, where $m(I - I^*) = 0$.

The sequence $\{x_{N_k}(t)\}$ converges uniformly in I_r for all values of r. That is to say, we have in I_r the relation

$$| x_{N_m}(t) - x_{N_n}(t) | \leqq \sum_{k=m}^{n-1} | x_{N_{k+1}}(t) - x_{N_k}(t) | < \sum_{k=m}^{n-1} \frac{1}{2^k} < \frac{1}{2^{m-1}}$$

$$(n > m > r).$$

Hence the sequence $\{x_{N_k}(t)\}$ converges in I^*; and by setting

$$x(t) = \begin{cases} \lim_{n\to\infty} x_{N_k}(t) & (t \in I^*), \\ 0 & (t \in I - I^*) \end{cases}$$

and bearing in mind the inequality

$$\int_{I_k(a)} | x_m(t) - x_{N_r}(t) |^p \, dt \leqq \int_{-\infty}^{\infty} | x_m(t) - x_{N_r}(t) |^p \, dt < \varepsilon$$

$$(m, N_r > N(\varepsilon)),$$

where $I_k(a)$ is the set of the elements $t \in I_k$ such that $-a \leqq t \leqq a$, and then passing to the limit in $I_k(a)$ (where the convergence of the sequence $\{x_{N_r}(t)\}$ is uniform), we obtain the inequality

$$\int_{I_k(a)} | x_m(t) - x(t) |^p \, dt < \varepsilon.$$

As $a \to \infty$ the left-hand side does not decrease. Therefore,

$$\int_{I_k} | x_m(t) - x(t) |^p \, dt \leqq \varepsilon,$$

where $m > N(\varepsilon)$ and k is arbitrary. Since $I_k \subset I_{k+1}$, we find as a result of this discussion that if $k \to \infty$ the integral

$$\int_{-\infty}^{\infty} | x_m(t) - x(t) |^p \, dt$$

exists and does not exceed ε, where $m > N(\varepsilon)$.

Whence it follows that $x_m - x \in L^p$, and consequently $x \in L^p$, which proves our theorem.

It will be noted that, in the course of our proof, the following was shown: If the sequence $\{x_n(t)\}_1^\infty$, in its normalized form, converges to $x(t)$ in L^p, i.e. if

$$\left\{\int_{-\infty}^{\infty} |x(t) - x_m(t)|^p \, dt\right\}^{\frac{1}{p}} \to 0,$$

then this sequence contains a subsequence $\{x_{N_k}(t)\}_1^\infty$ which converges to $x(t)$ almost everywhere.

26. Examples of Complete Orthonormalized Systems in L^2.

The functions

$$(1) \qquad \frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}}\cos t, \quad \frac{1}{\sqrt{\pi}}\sin t, \quad \frac{1}{\sqrt{\pi}}\cos 2t, \quad \frac{1}{\sqrt{\pi}}\sin 2t, \ldots$$

form a complete orthonormalized system in L^2 with respect to an interval of length 2π. In fact, it can be verified immediately that the system is orthogonal and normalized:

$$\int_0^{2\pi} \cos kt \sin mt \, dt = 0 \qquad (k = 0, 1, 2, \ldots;\ m = 1, 2, 3, \ldots);$$

$$\int_0^{2\pi} \sin kt \sin mt \, dt = \begin{cases} 0 & (k \neq m), \\ \pi & (k = m) \end{cases} \qquad (k, m = 1, 2, 3, \ldots);$$

$$\int_0^{2\pi} \cos kt \cos mt \, dt = \begin{cases} 0 & (k \neq m), \\ \pi & (k = m \neq 0), \\ 2\pi & (k = m = 0) \end{cases} \qquad (k, m = 0, 1, 2, \ldots).$$

The completeness of the system follows from the discussion in Section 18, since, as we have shown in Section 24, the totality of all trigonometric functions is a set lying dense in L^2.

Historically, the trigonometric system (1) was the first to appear in analysis. Fourier used trigonometric series for the first time to solve a problem in mathematical physics, and the reader knows from the elements of analysis that one designates as Fourier series (in the narrower sense) just those trigonometric series which can be associated with any integrable function, so that we have

$$f(t) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt),$$

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(t) \cos kt \, dt,$$

$$(k = 0, 1, 2, \ldots).$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(t) \sin kt \, dt.$$

From Section 18, formula (7), and

$$f(t) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt),$$

$$g(t) \sim \frac{c_0}{2} + \sum_{k=1}^{\infty} (c_k \cos kt + d_k \sin kt)$$

it follows that, for $f(t)$, $g(t) \in L^2$, we have the equation

$$\frac{1}{\pi} \int_0^{2\pi} f(t)\, \overline{g(t)}\, dt = \frac{a_0 \bar{c}_0}{2} + \sum_{k=1}^{\infty} (a_k \bar{c}_k + b_k \bar{d}_k),$$

(which is the generalized Parseval equation in the narrower sense). In particular, when $g(t) = f(t)$, we obtain the equation

$$\frac{1}{\pi} \int_0^{2\pi} |f(t)|^2\, dt = \frac{|a_0|^2}{2} + \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2).$$

Furthermore, as an immediate consequence of the completeness of the space L^2, we have the following theorem, known as the theorem of Fischer-Riesz [16]: *If the series*

$$\frac{|a_0|^2}{2} + \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2)$$

converges, then there exists a function $f(t) \in L^2$ having a_k, b_k $(k = 0, 1, 2, \ldots)$ as its Fourier coefficients, i.e. the equations

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt\, dt, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt\, dt \qquad (k = 0, 1, 2, \ldots)$$

are valid. This function is uniquely defined, except for a set of measure zero.

In fact, the sequence of functions

$$s_n(t) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt) \qquad (n = 1, 2, 3, \ldots)$$

converges by Cauchy's test, since

$$\frac{1}{\pi} \int_0^{2\pi} |s_m(t) - s_n(t)|^2\, dt = \sum_{k=m+1}^{n} (|a_k|^2 + |b_k|^2).$$

Hence, because of the completeness of the space L^2, there exists a function $f(t)$ for which the relation

$$\lim_{n\to\infty}\int_0^{2\pi}|f(t)-s_n(t)|^2\,dt=0$$

is fulfilled. Obviously, this function has the Fourier coefficients a_k, b_k $(k=1, 2, \ldots)$.

The functions

$$\frac{1}{\sqrt{\pi}},\quad \sqrt{\frac{2}{\pi}}\cos t,\quad \sqrt{\frac{2}{\pi}}\cos 2t,\ldots$$

form a complete orthonormalized system in the space L^2 with respect to the interval $[0, \pi]$.

It is required only to verify the completeness of this system. Let us assume that there exists a function $f(t)\in L^2$ such that

$$\int_0^{\pi}|f(t)|^2\,dt=1,\qquad \int_0^{\pi}f(t)\cos kt\,dt=0\qquad (k=0, 1, 2, \ldots)$$

is satisfied. If we extend this function on the interval $[-\pi, 0]$ with the aid of the equation

$$f(-t)=f(t)$$

we shall obtain

$$\int_{\pi}^{\pi}|f(t)|^2\,dt=2,\qquad \int_{-\pi}^{\pi}f(t)\cos kt\,dt=\int_{-\pi}^{\pi}f(t)\sin kt\,dt=0$$

$$(k=0, 1, 2, \ldots).$$

But that is impossible because of the completeness of the system previously studied.

We can prove similarly that the orthonormalized system

$$\sqrt{\frac{2}{\pi}}\sin t,\quad \sqrt{\frac{2}{\pi}}\sin 2t,\ldots$$

is complete in $L^2(0, \pi)$.

The Legendre polynomials

$$X_k(t)=\frac{1}{2^k\,k!}\frac{d^k(t^2-1)^k}{dt^k}\qquad (k=0, 1, 2, \ldots)$$

form a complete orthogonal system (but not a normalized system) in L^2 with respect to the interval $[-1, 1]$. The completeness thereof can be established in

the same way as was done at the beginning of this section for trigonometric functions. The orthogonality can be verified by integration by parts, viz.

$$2^{m+n}\, m!\, n! \int_{-1}^{1} X_m(t)\, X_n(t)\, dt =$$

$$= \frac{d^{n-1}(t^2-1)^n}{dt^{n-1}} \cdot \frac{d^m (t^2-1)^m}{dt^m} \Bigg|_{-1}^{1} - \int_{-1}^{1} \frac{d^{n-1}(t^2-1)^n}{dt^{n-1}} \frac{d^{m+1}(t^2-1)^m}{dt^{m+1}}\, dt =$$

$$= -\int_{-1}^{1} \frac{d^{n-1}(t^2-1)^n}{dt^{n-1}} \frac{d^{m+1}(t^2-1)^m}{dt^{m+1}}\, dt\,.$$

The expression

$$\frac{d^{n-k}(t^2-1)^n}{dt^{n-k}} \qquad (k = 1, 2, \ldots, n)$$

is clearly zero when $t = \pm 1$; hence, if $n > m$ and we continue the process of integration by parts, we finally obtain the expression

$$\pm \int_{-1}^{1} \frac{d^{n-m-1}(t^2-1)^n}{dt^{n-m-1}} \frac{d^{2m+1}(t^2-1)^m}{dt^{2m+1}}\, dt,$$

which is equal to zero.

27. Müntz's Theorem [17]. We will try to find a condition which must be satisfied by a sequence of distinct real numbers

$$p_1, p_2, p_3, \ldots \qquad \left(p_i > -\frac{1}{2}, \lim_{i\to\infty} p_i = \infty\right)$$

in order that the set of functions of the form

$$a_1 t^{p_1} + a_2 t^{p_2} + \ldots + a_n t^{p_n},$$

where $a_1, a_2, \ldots, a_n$ are arbitrary numbers and $n = 1, 2, \ldots$, shall fill densely L^2 or C with respect to the interval $[0, 1]$ or, what amounts to the same thing, so that the system of functions

$$t^{p_1}, t^{p_2}, t^{p_3}, \ldots$$

is closed in the space L^2 or C [18]. It is clear that in the case of the space C we must omit those values t^{p_i} for which $p_i < 0$.

Let us begin with the condition already found for the closure in L^2.

Since we have for the functions $f(t)$ of the space C:

$$\left\{\int_0^1 \left| f(t) - \sum_{i=1}^{n} a_i t^{p_i} \right|^2 dt \right\}^{\frac{1}{2}} \leqq \max_{0 \leqq t \leqq 1} \left| f(t) - \sum_{i=1}^{n} a_i t^{p_i} \right|$$

it follows that the condition which is necessary for the closure in L^2 is necessary also for the closure in C.

As the set of polynomials fills L^2 densely, it is necessary and sufficient for the closure of the system $\{t^{p_i}\}$ in L^2 that, corresponding to any integer $q \geq 0$ and any $\varepsilon > 0$, there exist an expression

$$\sum_{i=1}^{n} \lambda_i t^{p_i}$$

which satisfies the inequality

$$\int_0^1 \left| t^q - \sum_{r=1}^{n} \lambda_i t^{p_i} \right|^2 dt < \varepsilon .$$

In Section 14, we saw that the necessary and sufficient condition for this was that the relation

$$\lim_{n \to \infty} \prod_{i=1}^{n}{}' \left(\frac{q - p_i}{q + p_i - 1} \right)^2 = 0$$

be valid for any integer $q \geq 0$. (We may assume here that $q \neq p_i$). Our condition can be rewritten in the form

$$\lim_{n \to \infty} \prod_{i=1}^{n}{}' \left| \frac{1 - \dfrac{q}{p_i}}{1 + \dfrac{1+q}{p_i}} \right|^2 = 0 . \tag{1}$$

where the prime denotes the omission of one of the quantities $p_i = 0$.

Since by hypothesis $p_i \to \infty$, (1) will be true if and only if the series

$$\sum_{i=1}^{\infty}{}' \frac{1}{p_i}$$

diverges. Each of the products

$$\prod_{i=1}^{n}{}' \left(1 - \frac{q}{p_i}\right), \qquad \prod_{i=1}^{n}{}' \left(1 + \frac{q+1}{p_i}\right)$$

has of course a finite limit, different from zero, if the foregoing series converges as $n \to \infty$. If, however,

$$\sum_{i=1}^{\infty}{}' \frac{1}{p_i} = \infty ,$$

the first of these products approaches zero [15] and the second approaches infinity. Consequently, the system $\{t^{p_i}\}$ is closed in L^2 if and only if the series

$$\sum_{i=1}^{\infty}{}' \frac{1}{p_i}$$

diverges.

We now come to the space C. If we assume that n is a positive integer and if we eliminate from the system $\{p_i\}$ all negative numbers and zero, we find that, for $0 \leqq t \leqq 1$, the following relations hold:

$$(2) \qquad t^n - \sum_{i=1}^{m} \lambda_i t^{p_i} = n \int_0^t \left\{ \tau^{n-1} - \sum_{i=1}^{m} \mu_i \tau^{p_i - 1} \right\} dt \Big| \leqq$$

$$\leqq n \int_0^1 \left| \tau^{n-1} - \sum_{i=1}^{m} \mu_i \tau^{p_i-1} \right| dt \leqq$$

$$\leqq n \int_0^1 \left| \tau^{n-1} - \sum_{i=1}^{m} \mu_i \tau^{p_i-1} \right| dt \leqq$$

$$\leqq n \sqrt{\int_0^1 \left| \tau^{n-1} - \sum_{i=1}^{m} \mu_i \tau^{p_i-1} \right|^2 dt}.$$

Since we have

$$p_i \to \infty, \qquad \sum{}' \frac{1}{p_i} = \infty$$

whenever

$$p_i - 1 \to \infty, \qquad \sum{}' \frac{1}{p_i - 1} = \infty$$

it follows that the right-hand side of (2) can be made as small as we please by taking m large enough, and choosing μ_i appropriately.

Hence the quantities

$$\max_{0 \leqq t \leqq 1} \left| t^n - \sum_{1}^{m} \lambda_i t^{p_i} \right|$$

can be made as small as we please, for every integer $n > 0$.

[15] Translator's note: *I.e.* when $q > 0$, otherwise it approaches unity.

Therefore, if the number zero be also included among the characteristic numbers p_i, then the system of functions

$$\{t^{p_i}\} \quad (p_i \geqq 0, \lim_{i \to \infty} p_i = \infty),$$

for which

$$\sum{}' \frac{1}{p_i} = \infty$$

will be closed in C.

These two conditions, however, are not only sufficient, but also necessary. That follows from the fact that the condition

$$\sum{}' \frac{1}{p_i} = \infty ,$$

which, as we have shown, is necessary for the closure in L^2, is also sufficient. As to the condition that we should have , for a certain j,

$$p_j = 0 ,$$

such condition is evidently necessary for the closure in L^2. Indeed, if we assume every $p_i > 0$, then we shall find that all the expressions

$$\sum_{i=1}^{m} \lambda_i t^{p_i}$$

are zero when $t = 0$. Consequently, the error in the approximation of the function $x(t) = 1$ would be always ≥ 1, which is inconsistent with the closure of the system [19].

It will be observed that the conditions thus found are fulfilled if, for example, the sequence p_i is composed of prime numbers and zero.

28. The Concept of the Linear Functional. Let E be a certain linear normalized space and let G be a certain subspace therein. By a *linear functional* in G is meant a certain function $\varphi = \varphi(x)$ (defined for every $x \in G$) whose values are numbers and which satisfies the following conditions:

a) for any numbers α, β and any $x, y \in G$, we have

$$\varphi(\alpha x + \beta y) = \alpha\varphi(x) + \beta\varphi(y),$$

b) there exists a constant M such that

$$|\varphi(x)| \leqq M ,$$

if $||x|| < 1$ and $x \in G$.

The quantity

$$\sup_{\substack{x \in G \\ ||x|| \leqq 1}} |\varphi(x)| ,$$

which, according to b), does not exceed M, is called the *norm of the functional* φ and is denoted by the symbol $||\varphi||_G$; if $G = E$, we write simply $||\varphi||$.

If $x \in G$, then, from the definition of the norm, we have

$$\varphi\left(\frac{x}{\|x\|}\right) \leqq \|\varphi\|_G$$

or

$$|\varphi(x)| \leqq \|x\| \cdot \|\varphi\|_G \qquad (x \in G).$$

Whence the continuity of the linear functional follows: If $x_n \in G$ and $x_n \to x \in G$, then we have

$$\varphi(x_n) \to \varphi(x) \qquad (n \to \infty).$$

Suppose G_2 and $G_1 \subset G_2$ are two linear subspaces of E, and suppose $\varphi_1(x)$ and $\varphi_2(x)$ are linear functionals in G_1 and G_2. If $\varphi_1(x) = \varphi_2(x)$ for an arbitrary $x \in G_1$, then $\varphi_2(x)$ is called an extension of the functional $\varphi_1(x)$ on the subspace G_2.

In textbooks on functional analysis [20], there will be found a proof of the important theorem of Hahn-Banach: *Every linear functional $\varphi(x)$ which is defined on $G \subset E$ can be extended throughout the space E without increasing the norm. In other words, there exists in E a linear functional $\psi(x)$ such that*

$$\psi(x) = \varphi(x) \qquad (x \in G)$$

and

$$\|\psi\| = \|\psi\|_E = \|\varphi\|_G.$$

29. F. Riesz's Theorem [21]: *Every linear functional in a complete Hilbert space is of the form $\varphi(x) = (x, f)$, where f is a certain element belonging to H and uniquely determined by the functional φ; here, $\|\varphi\| = \|f\|$.*

To prove this, let us consider the set G of all those elements $y \in H$ which satisfy the equation

$$\varphi(y) = 0 \,.$$

Because the functional is linear, G is a certain subspace in H. This subspace is closed, since the functional is continuous. If $G = H$, the functional f is zero everywhere. The proposition holds if we put $f = 0$. Let us assume that $G \neq H$ and that f^* is an element of H for which

$$\inf_{y \in G} \|f^* - y\| = d > 0 \,.$$

According to Section 15, G contains one (and only one) element y^* such that

$$\|f^* - y^*\| = d.$$

Let us set

$$f_0 = \frac{1}{d}(f^* - y^*).$$

Then the equations

$$\|f_0\| = 1$$

and

$$(f_0, y) = 0$$

are valid for any value of $y \in G$. Let us set, also,

$$f = \varphi(f_0) f_0.$$

If x is any element of H, then the element

$$x - \frac{\varphi(x)}{\varphi(f_0)} f_0$$

causes φ to vanish and hence belongs to G. Thus

$$\left(f_0, x - \frac{\varphi(x)}{\varphi(f_0)} f_0\right) = 0.$$

Therefore

$$(f_0, x) = \left(\frac{\overline{\varphi(x)}}{\overline{\varphi(f_0)}}\right)(f_0, f_0)$$

or

$$\overline{\varphi(x)} = \overline{\varphi(f_0)}\,(f_0, x),$$

i.e.

$$\varphi(x) = (x, f).$$

We are going to prove that the element f is uniquely determined. Suppose that, for an arbitrary value of $x \in H$, we have

$$\varphi(x) = (x, f') = (x, f'') ,$$

where $f' \neq f''$. If we set $x = f' - f''$, we get

$$(f' - f'', f') = (f' - f'', f'').$$

Whence it follows that

$$(f' - f'', f' - f'') = 0,$$

which contradicts our hypothesis.

There remains for us to show that

$$\| \varphi \| = \| f \|.$$

Now, by the rule for the inner product, we obtain

$$| \varphi(x) | \leqq \| f \| \, \| x \| .$$

Hence

$$\| \varphi \| \leqq \| f \|.$$

But, on the other hand,

$$| \varphi(f) | = \| f \|^2,$$

and consequently

$$\| \varphi \| \geqq \| f \| .$$

This completes the proof of Riesz's theorem.

By virtue of this theorem, a linear functional in $L^2(a, b)$ is of the form

$$\varphi(x) = \int_a^b x(t) f(t)\, dt,$$

where $f = f(t) \in L^2(a, b)$ and

$$\| \varphi \| = \| f \|.$$

It can be shown [22] that a linear functional in $L^p(a, b)$ $(p > 1)$ has the same form, except that $f = f(t) \in L^q(a, b)$ and

$$\| \varphi \| = \| f \|_q \quad \left(\frac{1}{p} + \frac{1}{q} = 1 \right).$$

Furthermore, a linear functional in L has the same form, but $f = f(t)$ is now a measurable function such that

$$\text{true max} \, | f(t) | = \| \varphi \|.$$

Therefore, we have

$$\text{true max} \, | f(t) | = \inf_{I^*} \sup_{t \in I - I^*} | f(t) |,$$

where I is the interval (a, b) and I^* is a set therein of measure zero. Thus, the true maximum is the greatest lower bound of all the least upper bounds of $f(t)$, and is obtained after excluding from the interval (a, b) all possible sets of measure zero.

Finally, a linear functional in the space $C(0, 1)$ has the form [22]

$$\varphi(x) = \int_0^1 x(t)\, d\sigma(t),$$

where $\sigma(t)$ is a function of bounded variation. It is understood that the integral is a Stieltjes integral and that we are to put

$$\| \varphi \| = \operatorname*{Var}_{0 \leq t \leq 1} \sigma(t).$$

30. A Criterion for the Closure of a Set of Vectors in Linear Normalized Spaces [23]. A set $\mathfrak{M}$ of vectors of the linear normalized space E is said to be closed in E if every vector $x \in E$ can be approximated to any degree of accuracy by means of an expression of the form

$$a_1 y_1 + a_2 y_2 + \ldots + a_n y_n, \tag{1}$$

where $y_i \in \mathfrak{M}$.

The necessary and sufficient condition that a set $\mathfrak{M}$ in E be a closed set is, that every linear functional $\varphi(x)$ in E which vanishes for any vector $y \in \mathfrak{M}$ be identically zero.

Proof: Let $\mathfrak{M}$ be a closed set in E, and let the linear functional $\varphi(x)$ be zero on every vector $y \in \mathfrak{M}$. Then, if x is any vector of E, and if ε is any positive number,

we can find an expression of the form (1) which satisfies the inequality

$$\| x - a_1 y_1 - a_2 y_2 - \ldots - a_n y_n \| \leqq \varepsilon \ .$$

Since

$$| \varphi (x - a_1 y_1 - a_2 y_2 - \ldots - a_n y_n) | \leqq \| \varphi \| \, \| x - a_1 y_1 - \ldots - a_n y_n \|$$

we have

$$| \varphi (x) | \leqq \| \varphi \| \varepsilon \ ;$$

and consequently $\varphi(x) = 0$, inasmuch as ε was chosen arbitrary. The necessity for the condition is thus established.

In order to prove the sufficiency of the condition, let us assume that there exists a vector $x_0 \in E$ such that

$$\inf_{y_i \in M} \| x_0 - a_1 y_1 - a_2 y_2 - \ldots - a_n y_n \| = d > 0 \ .$$

Let G denote a linear subspace of E whose elements are vectors of the form

$$a_1 y_1 + a_2 y_2 + \ldots + a_n y_n + \beta x_0 \qquad (y_i \in \mathfrak{M}) \ . \tag{2}$$

Let the equation

$$\varphi (x) = \beta$$

define a linear functional in G, for every x of the form (2). This definition is permissible, since the vector $x \in G$ cannot have two representations of the form (2), each with a different β.

Since

$$\| a_1 y_1 + a_2 y_2 + \ldots + a_n y_n + \beta x_0 \| \geqq | \beta | \, d$$

it follows that the norm of the functional $\varphi(x)$ in G is not greater than $\frac{1}{d}$. On the other hand, we have $\lim_{k \to \infty} \| x_0 - z_k \| = d$

where

$$z_k = \sum_{i=1}^{n_k} a_i^{(k)} y_i \qquad (y_i \in \mathfrak{M}) \ .$$

Hence, from

$$1 = | \varphi (x_0 - z_k) | \leqq \| \varphi \|_G \, \| x_0 - z_k \|$$

we obtain the inequality

$$\| \varphi \|_G \geqq \frac{1}{d} \ .$$

Consequently, $\| \varphi \|_G = \frac{1}{d}$.

If we extend the functional φ throughout the entire space E by means of the Hahn-Banach theorem, we obtain a functional in E. This functional is not identically zero, but it does vanish on all the vectors $y \in \mathfrak{M}$. Thus the sufficiency of the condition is established.

CHAPTER II

P. L. TCHEBYSHEFF'S DOMAIN OF IDEAS

31. Statement of the Problem. Given a closed (finite or infinite) interval $[a, b]$ on the number axis, and two real-valued functions $f(x)$, $s(x)$ continuous in $[a, b]$. Let us form the expression

$$\text{(T)} \qquad Q(x) = s(x)\,\frac{q_0x^n + q_1x^{n-1} + \cdots + q_n}{p_0x^m + p_1x^{m-1} + \cdots + p_m},$$

where m, n are given; and let us propose the problem of determining real parameters $p_0, p_1, \ldots, p_m$; $q_0, q_1, \ldots, q_n$ so that the deviation of $Q(x)$ from $f(x)$, namely

$$H_Q = \max_{a \leqq x \leqq b} |f(x) - Q(x)|$$

shall be a minimum.

In the special case where $s(x) = 1$, $m = 0$ and the interval $[a, b]$ is finite, the proposed problem is reduced to the problem of finding the best approximation of a given function in the space C through polynomials of the n-th degree.

The problem of finding the minimum of H_Q leads to the inquiry as to the best approximation of a given continuous periodic function in terms of trigonometric sums of the n-th order;. (Cf. Sections 4 and 40.)

With respect to the function $s(x)$, we will assume that it is different from zero inside the interval (a, b). If the interval $[a, b]$ is finite, the same thing shall be true of the end points of the interval. If the interval $[a, b]$ is infinite, we will not exclude the possibility of the function $s(x)$ vanishing at infinity. Instead, we will require in such case that, for a certain integer $k \geq 0$, the product

$$s(x)\,x^k$$

shall tend towards a finite limit different from zero as $x \to \infty$. Let us assume that $m = n - k$. In addition, if the interval is the entire number axis, we will assume that

$$\lim_{x \to \infty} s(x)\,x^k = \lim_{x \to -\infty} s(x)\,x^k$$

(whence it follows that k in this case is surely an even number), and will consider only such functions $f(x)$ for which

$$\lim_{x \to \infty} f(x) = \lim_{x \to -\infty} f(x)\,.$$

With this last condition, the space C_∞ becomes the subject of our study, and it is convenient to identify the points $\pm\infty$.

Moreover, since all real polynomials of odd degree have at least one real root, we may assume in the case where $[a, b]$ is the entire axis that the number m is even (otherwise we reduce the number of parameters p_i by one, at the outset).

From here on, in this discussion, we will assume that all the conditions enumerated above are fulfilled; they will not be mentioned specifically on each occasion.

32. A Generalization of the Theorem of de la Vallée-Poussin [24]: *If the polynomials*

$$A(x) = a_0 x^{m-\mu} + \ldots + a_{m-\mu}, \qquad B(x) = b_0 x^{n-\nu} + \ldots + b_{n-\nu},$$

where

$$0 \leqq \mu \leqq m, \quad 0 \leqq \nu \leqq n, \quad a_0 \neq 0$$

have no common divisor and the expression

$$R(x) = s(x) \frac{B(x)}{A(x)}$$

remains finite in the interval $[a, b]$, *and if the difference*

$$f(x) - R(x)$$

at the consecutive points

$$x_1 < x_2 < \ldots < x_N \tag{1}$$

of the interval $[a, b]$ *assumes values*

$$\lambda_1, -\lambda_2, \ldots, (-1)^{N-1}\lambda_N$$

different from zero and alternately plus and minus (in that connection, we may select any $\lambda_i > 0$), *where*[1] $N = m + n - d + 2$ *with* $d = \min(\mu, \nu)$, *then the inequality*

$$H_Q \geqq \min\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$$

is valid for every function $Q(x)$ *of the form* (T).

The same inequality is still valid if $R(x) \equiv 0$ and $N = n + 2$, i.e., $d = m$.

For the proof, let us assume that some function $Q(x)$ of the form (T) satisfies the inequality

$$H_Q < \min\{\lambda_1, \lambda_2, \ldots, \lambda_N\}.$$

We now form the difference

$$\Delta(x) = Q(x) - R(x) = [f(x) - R(x)] - [f(x) - Q(x)].$$

[1] If the interval $[a, b]$ is the entire number axis, the number N can only be even because of our hypothesis in such case. Hence in this case we may speak of the consecutive points (1) of the semi-closed interval $[-\infty, \infty)$, which is the same as saying that the points $\pm\infty$ can be regarded as one point.

Obviously, the numbers

$$\Delta(x_1), \Delta(x_2), \ldots, \Delta(x_N)$$

are different from zero and change signs alternately. But since $\Delta(x)$ is continuous in the interval $[a, b]$, it follows that $\Delta(x)$ has at least $N - 1 = m + n + 1 - d$ zeros in the open interval (a, b). But that is impossible, inasmuch as

$$\Delta(x) = s(x) \frac{V(x)}{U(x)},$$

where the second of the polynomials $U(x)$, $V(x)$ is of degree $\leqq m + n - d$.

Thus our theorem is established.

The significance of this theorem lies in the fact that it enables us to obtain a certain lower estimate for the deviation of the best approximation.

33. The Existence Theorem. We are going to show that *among the functions* (T) *there exists at least one function for which* H_Q *has a minimum.*

Let $H \geqq 0$ be the greatest lower bound of the set of all the quantities H_Q. Then, by definition, there exists an infinite sequence of functions $Q_i(x)$ $(i = 1, 2, \ldots)$ such that

$$H_{Q_i} \to H.$$

Let us now so normalize the functions $Q_i(x)$, that we have

$$p_{i0}^2 + p_{i1}^2 + \ldots + p_{im}^2 = 1 \qquad (i = 1, 2, 3, \ldots)$$

and prove that for this normalization all the quantities q_{ij} are bounded.

Let

$$H_{Q_i} < G \qquad (i = 1, 2, 3, \ldots),$$

and let $\xi_1, \xi_2, \ldots, \xi_{n+1}$ represent certain fixed, distinct points inside the interval (a, b); then we shall have

$$\left| s(\xi) \frac{q_{i0}\xi^n + q_{i1}\xi^{n-1} + \ldots + q_{in}}{p_{i0}\xi^m + p_{i1}\xi^{m-1} + \ldots + p_{im}} \right| \leqq G + \max_{a \leqq x \leqq b} |f(x)|,$$

where ξ is any one of the points ξ_i, and hence

$$|q_{i0}\xi^n + q_{i1}\xi^{n-1} + \ldots + q_{in}| < K.$$

If the values of the polynomials

$$q_{i0}x^n + q_{i1}x^{n-1} + \ldots + q_{in}$$

are bounded at $n + 1$ fixed points, then all the coefficients of these polynomials are also bounded.

We now make use of the fact that the numbers

$$p_{ij}, q_{ik} \qquad \left(i = 1, 2, \ldots; \begin{matrix} j = 0, 1, \ldots, m \\ k = 0, 1, \ldots, n \end{matrix} \right)$$

are bounded, and we select from the sequence $\{Q_i(x)\}$ a certain subsequence (which will likewise be denoted by $\{Q_i(x)\}$) such that the limits

$$\lim_{i\to\infty} p_{ij} = a_j, \quad \lim_{i\to\infty} q_{ik} = b_k$$

exist. Thus we obtain a function of the form (T):

$$P(x) = s(x)\,\frac{b_0x^n + b_1x^{n-1} + \ldots + b_n}{a_0x^m + a_1x^{m-1} + \ldots + a_m};$$

we will show that the function $P(x)$ satisfies the relation

$$H_P \equiv \max_{a\leqq x\leqq b} |f(x) - P(x)| = H.$$

$P(x)$ can become infinite at only a finite number of points. At those points $\tilde{x}$ of the interval $[a, b]$ which are different from the aforementioned points and, in the case of an infinite interval $[a, b]$, are different from the points at infinity as well, we have obviously

$$P(\tilde{x}) = \lim_{i\to\infty} Q_i(\tilde{x})|;$$

hence

$$|P(\tilde{x})| \leqq |f(\tilde{x})| + |f(\tilde{x}) - Q_i(\tilde{x})| + |Q_i(\tilde{x}) - P(\tilde{x})| \leqq \max_{a\leqq x\leqq b} |f(\tilde{x})| + H_{Q_i} + \varepsilon_i,$$

where $\varepsilon_i \to 0$ as $i \to \infty$.

Consequently, at every point in $[a, b]$, except perhaps for a finite number of points (and hence almost everywhere) we have

$$|P(x)| < M = \max_{a\leqq x\leqq b} |f(x)| + G.$$

Therefore, $Q_i(x)$ converges uniformly to $P(x)$ in $[a, b]$ if the interval $[a, b]$ is finite; and, if the interval $[a, b]$ is infinite, $Q_i(x)$ converges uniformly to $P(x)$ in every finite closed subinterval $[a, \beta]$.

In the latter case, we have

$$\max_{a\leqq x\leqq\beta} |f - P| \leqq \max_{a\leqq x\leqq\beta} |f - Q_i| + \max_{a\leqq x\leqq\beta} |P - Q_i|;$$

whence

$$\max_{a\leqq x\leqq\beta} |f - P| \leqq H$$

and accordingly,

$$H_P \leqq H.$$

The same inequality results in the case of a finite interval $[a, b]$. However, since

$$H_P \geqq H,$$

it follows that

$$H_P = H,$$

and our theorem is proved.

34. Tchebysheff's Theorem [25] says: *The function $P(x)$, which deviates less from the function $f(x)$ than does any other function of the form* (T) *in* $[a, b]$ (Section 31), *is uniquely determined if we consider two fractions as identical when they coincide after being reduced to their lowest terms.*

This function is completely characterized by the following property: If the function can be expressed in the form

$$P(x) = s(x)\,\frac{b_0 x^{n-\nu} + b_1 x^{n-\nu-1} + \ldots + b_{n-\nu}}{a_0 x^{m-\mu} + a_1 x^{m-\mu-1} + \ldots + a_{m-\mu}} = s(x)\,\frac{B(x)}{A(x)}$$

where $0 \leq \mu \leq m$, $0 \leq \nu \leq n$, $a_0 \neq 0$ *and the fraction* $\dfrac{B(x)}{A(x)}$ *is irreducible, then the number* N *of consecutive points* [2] *of the interval* $[a, b]$ *at which the difference* $f(x) - P(x)$, *with alternate change of sign, takes on the value* H_P, *is not less than* $m + n + 2 - d$, *where* $d = \min(\mu, \nu)$; *in case* $P(x) \equiv 0$, *then* $N \geq n + 2$.

Proof: If we set

$$|\lambda_k| = H_P \qquad (k = 1, 2, \ldots, N)$$

in the generalized theorem of de la Vallée-Poussin (Section 32), it follows from said theorem that, among all functions of the form (T), the function $P(x)$ deviates the least from $f(x)$ when $N \geq m + n - d + 2$. Let us now assume that the number of "deviation points" in the interval $[a, b]$ is

$$N' \leqq m + n - d + 1\,;$$

we will prove that in such case $P(x)$ is not the function, of the form (T), which deviates the least from $f(x)$.

By hypothesis, it follows that the interval $[a, b]$ can be subdivided into N' partial intervals

$$[a, \xi_1],\ [\xi_1, \xi_2],\ \ldots,\ [\xi_{N'-1}, b] \tag{1}$$

in such a way that in each of them in turn one of the following two inequalities

$$\begin{cases} -H_P \leqq f(x) - P(x) < H_P - \alpha, \\ -H_P + \alpha < f(x) - P(x) \leqq H_P \end{cases} \tag{2}$$

is satisfied. Here α is a certain positive number. Let us now consider the function

$$\Phi(x) = (x - \xi_1)(x - \xi_2) \ldots (x - \xi_{N'-1}).$$

As the polynomials $A(x)$, $B(x)$ have no common divisor, we can find polynomials $\varphi(x)$, $\psi(x)$ of degrees m, n, respectively, such that

$$\Phi(x) = A(x)\,\psi(x) - B(x)\,\varphi(x).$$

---

[2] If the interval $[a, b]$ is the entire number axis, the points $\pm\infty$ will be considered as one point. If nothing to the contrary is said, we shall consider, in this case, the semi-closed interval $[-\infty, \infty)$ in place of $[-\infty, \infty]$. Cf. footnote [1], Chapter II.

Let us now construct a function of the form (T) according to the formula

$$Q(x) = s(x)\frac{B(x)\,\Omega(x) - \omega\psi(x)}{A(x)\,\Omega(x) - \omega\varphi(x)},$$

where ω is a real parameter and Ω a polynomial of degree $\leq d$ which is to be appropriately defined.

Now we have

$$f(x) - Q(x) = f(x) - P(x) + s(x)\frac{\omega\Phi(x)}{A(x)\,\{A(x)\,\Omega(x) - \omega\varphi(x)\}}.$$

The numerator of the last term on the right-hand side changes sign at $N' - 1$ points

$$\xi_1 < \xi_2 < \ldots < \xi_{N'-1}.$$

If we prove that the polynomial $\Omega(x)$ can be chosen so that, for all sufficiently small values of $|\omega|$, the denominator

$$A(x)\,\{A(x)\,\Omega(x) - \omega\varphi(x)\}$$

is positive in the interval $[a, b]$ and the fraction

$$s(x)\frac{\Phi(x)}{A(x)\,\{A(x)\,\Omega(x) - \omega\varphi(x)\}}$$

is bounded, then our proposition will be an immediate consequence of the inequalities (2), which are satisfied alternately in the intervals (1). Let us show, then, how to choose the polynomial $\Omega(x)$.

If the interval $[a, b]$ is finite, we can choose $\Omega(x) \equiv 1$.

If the interval $[a, b]$ is infinite, then, by hypothesis, $m = n - k$ (see end of Section 31). Hence in this case $\mu \leq \nu$ and consequently $\mu = d$. It follows that we can take for $\Omega(x)$ any polynomial of precisely the degree μ, and positive in the interval $[a, b]$. In this connection, it must be remembered that if the interval is the entire number axis and hence m is an even number, then μ is also an even number.

Thus we have demonstrated the characteristic property of the function which yields the minimum deviation.

The only thing that remains to be proved is the uniqueness of this function. The proof is accomplished by the indirect method. Let us assume that, besides the function $P(x)$, there exists another function $Q(x)$ of the form (T) such that

$$H_Q = H_P = H$$

Suppose the numbers N', μ', ν', d' play the same role for $Q(x)$ that the numbers N, μ, ν, d do for $P(x)$, where

$$N' \geqq m + n + 2 - d',\ N \geqq m + n + 2 - d\ .$$

Without loss of generality, we may assume that $N' \geq N$. Let

$$\beta_1 < \beta_2 < \ldots < \beta_{N'}$$

be the deviation points for $Q(x)$. We now form the difference

$$\Delta(x) = P(x) - Q(x) = [f(x) - Q(x)] - [f(x) - P(x)],$$

which may vanish at certain points β_j. If $\Delta(\beta_j) \neq 0$, it is easy to see that we then have

$$\text{sign}\,[f(\beta_j) - Q(\beta_j)] = \text{sign}\,\Delta(\beta_j). \tag{3}$$

Suppose, for example, that

$$\Delta(\beta_{i-1}) \neq 0,\ \Delta(\beta_i) = \Delta(\beta_{i+1}) = \ldots = \Delta(\beta_{i+k}) = 0,\ \Delta(\beta_{i+k+1}) \neq 0. \tag{4}$$

Inasmuch as

$$(-1)^{i-1}[f(\beta_{i-1}) - Q(\beta_{i-1})],\ (-1)^{i+k+1}[f(\beta_{i+k+1}) - Q(\beta_{i+k+1})]$$

have one and the same sign, the numbers

$$\Delta(\beta_{i-1}),\ (-1)^k\Delta(\beta_{i+k+1})$$

have like signs, by virtue of (3). Hence the number of zeros of the difference $\Delta(x)$ in the interval $[\beta_{i-1}, \beta_{i+k+1}]$ is even or odd according as the number k is even or odd. By reason of (4), this number of zeros is at least $k + 2$.

By repeating this reasoning we find that $\Delta(x)$ has at least $N' - 1$ zeros in the interval (a, b). But that is impossible, since

$$\Delta(x) = s(x)\,\frac{V(x)}{U(x)}$$

and the second of the polynomials $U(x)$, $V(x)$ is of degree

$$r = \begin{cases} \max(m + n - \mu' - \nu,\ m + n - \mu - \nu') \leq N' - 2, & \text{when } P \not\equiv 0,\ Q \not\equiv 0, \\ n - \nu' \leq N' - 2, & \text{when } P \equiv 0, \\ n - \nu \leq N - 2 \leq N' - 2, & \text{when } Q \equiv 0. \end{cases}$$

Thus the theorem is fully demonstrated.

35. A Special case of Tchebysheff's Theorem. If $s(x) \equiv 1$ and $m = 0$ and the interval $[a, b]$ is finite, we have the special case already mentioned in Section 31; this case is particularly important. In this case, we get the theorem: *Among all polynomials of the n-th degree in metric space C, the polynomial $P(x)$ which deviates the least from a given continuous function is uniquely characterized by the fact that in the interval $[a, b]$ the number of consecutive points at which the difference $f(x) - P(x)$ with alternate change in sign assumes the value* $\max_{a \leq x \leq b} |f(x) - P(x)|$, *is not less than* $n + 2$.

36. The Tchebysheff Polynomials of Least Deviation from Zero. As an example of Tchebysheff's theorem, let us consider the problem: Among all the polynomials of the n-th degree whose leading coefficient is unity, it is required to find a particular polynomial, the maximum of whose absolute value in the interval $(-1, 1)$

is less than that of any other of the said polynomials. In order to be in a position to utilize the general theorem, we seek a suitable polynomial of degree $n-1$ which deviates the least from the function $f(x) = x^n$. Hence the polynomial required by Tchebysheff's theorem, viz.

$$T_n(x) = x^n + p_1 x^{n-1} + \ldots + p_n$$

is completely characterized by the fact that the number of consecutive points of the interval $[-1, 1]$, such that $T_n(x)$ with alternate change in sign assumes the value $\max_{-1 \leq x \leq 1} | T_n(x) |$ at these points, is not less than $n+1$.

We will show that:

$$T_n(x) = \frac{1}{2^n}\{(x + \sqrt{x^2-1})^n + (x - \sqrt{x^2-1})^n\}.$$

Evidently, $T_n(x)$ is a polynomial of the n-th degree. That its leading coefficient is equal to unity follows from the formula

$$\lim_{x \to \infty} \frac{T_n(x)}{x^n} = \frac{1}{2^n} \lim_{x \to \infty} \left(1 + \sqrt{1 - \frac{1}{x^2}}\right)^n.$$

Setting $x = \cos \Theta (0 \leq \Theta \leq \pi)$, we obtain

$$T_n(x) = \frac{1}{2^{n-1}} \cos n\Theta.$$

Hence

$$\max_{-1 \leq x \leq 1} | T_n(x) | = \frac{1}{2^{n-1}}.$$

Let us consider the points

$$x_k = -\cos \frac{k\pi}{n} \qquad (k = 0, 1, 2, \ldots, n).$$

Since

$$T_n(x_k) = \frac{1}{2^{n-1}} \cos (n-k)\pi = \frac{(-1)^{n-k}}{2^{n-1}} \qquad (k = 0, 1, 2, \ldots, n)$$

it follows that $T_n(x)$, with alternate change in sign at the points

$$-1 = x_0 < x_1 < \ldots < x_{n-1} < x_n = 1$$

assumes its maximum value $\frac{1}{2^{n-1}}$ at those points and hence actually represents a solution of the given problem. $T_n(x)$ is called *Tchebysheff's polynomial of the n-th degree which deviates least from zero.*

37. A Further Example of P. Tchebysheff's Theorem. As a second example of the theorem of P. Tchebysheff in Section 35, we will consider the following

problem [26]: It is required to find the deviation of the best approximation of the function $\frac{1}{x-a}$ in the interval $[-1, 1]$ (by polynomials of the n-th degree), where $a > 1$.

Let us study the function

$$\Phi(x) = \frac{M}{2}\left\{v^n \frac{a-v}{1-av} + v^{-n}\frac{1-av}{a-v}\right\} \tag{1}$$

where

$$x = \frac{1}{2}\left(v + \frac{1}{v}\right), \qquad a = \frac{1}{2}\left(a + \frac{1}{a}\right), \qquad a = a - \sqrt{a^2-1} < 1.$$

As $\Phi(x)$ is a rational function of v and remains unaltered when v is substituted for $\frac{1}{v}$, it follows that $\Phi(x)$ is a rational function of x.

By observing that

$$\lim_{x\to a}(x-a)\,\Phi(x) = \frac{M}{2}\lim_{v\to a}\frac{(v-a)(av-1)}{2av}\,v^{-n}\,\frac{1-av}{a-v} = \frac{M(1-a^2)^2}{4a^{n+2}}$$

we conclude that, when

$$M = \frac{4a^{n+2}}{(1-a^2)^2},$$

the function $\Phi(x)$ has the form

$$\Phi(x) = \frac{1}{x-a} - P_n(x), \tag{2}$$

where $P_n(x)$ is a polynomial of the n-th degree with real coefficients.

If the point x runs through the interval $[-1, 1]$ from right to left, v will describe the upper half of the circle $|v| = 1$ in a positive direction. From the formula (1), it follows that, in the interval $[-1, 1]$, $\Phi(x)$ does not exceed M because, when $|v| = 1$, we have

$$\left|v^n \frac{a-v}{1-av}\right| = 1.$$

Moreover, it will be noted that $\Phi(x)$ assumes the value M if

$$\arg\left\{v^n \frac{a-v}{1-av}\right\} \equiv 0 \pmod{2\pi}$$

and the value $-M$ if

$$\arg\left\{v^n \frac{a-v}{1-av}\right\} \equiv \pi \pmod{2\pi}.$$

Since the function

$$v^n \frac{a-v}{1-av}$$

has an n-tuple zero at the origin of the circle $|v| < 1$, and a simple zero at the point α, it is evident that

$$\arg\left\{v^n \frac{\alpha - v}{1 - \alpha v}\right\}$$

increases from π to $(n + 2)\pi$ if the point v describes the upper half of the circle $|v| = 1$ in a positive direction.

Therefore the function $\Phi(x)$, with alternate change in sign, assumes the value M at $n + 2$ consecutive points of the interval $[-1, 1]$. Consequently, among all polynomials of the n-th degree, the polynomial $P_n(x)$, which is defined by the formulas (1) and (2), deviates the least from $\frac{1}{x - \alpha}$ in the interval $[-1, 1]$.

Thus the required deviation of the best approximation is

$$M = \frac{4\alpha^{n+2}}{(1 - \alpha^2)^2} = \frac{(\alpha - \sqrt{\alpha^2 - 1})^n}{\alpha^2 - 1}.$$

38. An Example for the Application of the General Theorem of de la Vallée-Poussin. We are now going to consider the same sort of problem as in the previous section, but, instead of the function $\frac{1}{x - \alpha}$, here we shall deal with the function

[27]

$$f(x) = \frac{A(\alpha^2 - 1)^2}{(x - \alpha)^2} + \frac{A'(\alpha^2 - 1)}{x - \alpha}$$

where we assume A, A' and $\alpha > 1$ are given real numbers. We retain the previous meaning of the quantities v and α and study the function

$$\Phi(x) = \frac{L}{2}\left\{v^n \left(\frac{\alpha - v}{1 - \alpha v}\right)^2 \frac{1 - \beta v}{v - \beta} + v^{-n}\left(\frac{1 - \alpha v}{\alpha - v}\right)^2 \frac{v - \beta}{1 - \beta v}\right\}.$$

If we set

$$L = \frac{A\sqrt{\alpha^2 - 1}}{(\alpha + \sqrt{\alpha^2 - 1})^n}\lambda, \quad \beta = \alpha\left\{1 - \left(\lambda - \frac{\alpha^2}{1 - \alpha^2}\right)^{-1}\right\}, \quad b = \frac{1}{2}\left(\beta + \frac{1}{\beta}\right)$$

$$\lambda = \frac{1}{2}\left\{n + \frac{2\alpha A - A'}{A\sqrt{\alpha^2 - 1}} + \sqrt{\left(n + \frac{2\alpha A - A'}{A\sqrt{\alpha^2 - 1}}\right)^2 + \frac{1}{\alpha^2 - 1}}\right\}$$

where the root is so defined that

$$\left(n + \frac{2\alpha A - A'}{A\sqrt{\alpha^2 - 1}}\right)\sqrt{\left(n + \frac{2\alpha A - A'}{A\sqrt{\alpha^2 - 1}}\right)^2 + \frac{1}{\alpha^2 - 1}} > 0,$$

we obtain, after some simplification:

$$\Phi(x) = \frac{A(\alpha^2 - 1)^2}{(x - \alpha)^2} + \frac{A'(\alpha^2 - 1)}{x - \alpha} - P_n(x) + \frac{B}{x - b},$$

where $P_n(x)$ is a polynomial of the n-th degree.

Therefore we find, in the same way as we did before, that the maximum value of the function in the interval $[-1, 1]$ is $|L|$, where $\Phi(x)$ assumes this value with alternate change in sign at $n+2$ consecutive points of the interval $[-1, 1]$.

On the other hand, it can be shown without difficulty that, when

$$n > \frac{2A\sqrt{a^2-1}+A'}{A\sqrt{a^2-1}}$$

we obtain the inequality

$$\max_{-1\leq x\leq 1}\left|\frac{B}{x-b}\right| < \mu\,|L|$$

with

$$\mu = \frac{1}{4}\frac{(a-\sqrt{a^2-1})^n}{a-1}\cdot\frac{1}{\left(n+\dfrac{2aA-A'}{A\sqrt{a^2-1}}\right)^2}.$$

Therefore, if n is chosen sufficiently large, the difference

$$\left\{\frac{A(a^2-1)^2}{(x-a)^2}+\frac{A'(a^2-1)}{x-a}\right\}-P_n(x)$$

assumes at $n+2$ consecutive points of the interval alternately positive and negative values and, in absolute magnitude, these values exceed

$$|L|\,(1-\mu)\,.$$

By reason of de la Vallée-Poussin's theorem [$s(x)\equiv 1$, $m=0$] the required deviation of the best approximation is not less than

$$|L|\,(1-\mu),$$

and on the other hand the said deviation is not greater than

$$|L|\,(1+\mu).$$

Denoting the required deviation by $\mathsf{E}_n[f]$, we obtain

$$\mathsf{E}_n[f] = \mathsf{E}_n\left[\frac{A(a^2-1)^2}{(x-a)^2}+\frac{A'(a^2-1)}{x-a}\right]=$$

$$=\frac{1}{2}\frac{|A|\sqrt{a^2-1}}{|(a+\sqrt{a^2-1})^n|}\left|n+\frac{2aA-A'}{A\sqrt{a^2-1}}+\sqrt{\left(n+\frac{2aA-A'}{A\sqrt{a^2-1}}\right)^2+\frac{1}{a^2-1}}\right|(1+\varepsilon_n)$$

with

$$|\varepsilon_n| < \frac{1}{4}\frac{(a-\sqrt{a^2-1})^n}{a-1}\cdot\frac{1}{\left(n+\dfrac{2aA-A'}{A\sqrt{a^2-1}}\right)^2},$$

and hence

$$\lim_{n \to \infty} \varepsilon_n = 0.$$

39. An Example for the Application of P. L. Tchebysheff's General Theorem. Elliptic functions are employed in the example to be considered in this section. First of all, we will formulate the problem: Among all rational fractions in which both numerator and denominator are of the m-th degree, it is required to find that particular fraction by means of which the function

$$\frac{1}{\sqrt{1 - k^2 x}}$$

can be approximated so that the maximum value of the relative error in the interval [0, 1] shall be the least; here the given number k shall satisfy the inequality

$$0 < k < 1 .$$

What confronts us here is obviously the general case of the problem in Section 31 for

$$n = m, \; f(x) = 1, \; s(x) = \sqrt{1 - k^2 x}$$

Let us recall certain properties of elliptic functions.

Jacobi's function $\xi = sn\,(u; k)$ is defined by the relation

$$u = \int_0^{\xi} \frac{dt}{\sqrt{(1 - t^2)\,(1 - k^2 t^2)}} ,$$

the function $\mathrm{dn}(u; k)$, by the equations

$$\mathrm{dn}(u; k) = \sqrt{1 - k^2\,\mathrm{sn}^2(u; k)}, \qquad \mathrm{dn}\,(0; k) = 1 .$$

If we put

$$K = \int_0^1 \frac{dt}{\sqrt{(1 - t^2)\,(1 - k^2 t^2)}}$$

and

$$x = \mathrm{sn}^2(u; k),$$

then x increases from 0 to 1 as u increases from 0 to K.

$2K$ is the period of the function

$$\mathrm{sn}^2\,(u + 2K; k) = \mathrm{sn}^2(u; k).$$

k is called the modulus of our function and K the complete elliptic integral of the first kind for the modulus k.

These are well-known facts. To the realm of elementary facts there belongs also the relation

(1) $$\operatorname{dn}\left(\frac{u}{M}\,;\,\lambda\right) = \operatorname{dn}(u;k)\prod_{r=1}^{m}\frac{1-k^2C_{2r-1}\operatorname{sn}^2(u;k)}{1-k^2C_{2r}\operatorname{sn}^2(u;k)},$$

which occurs in the so-called theory of transformations of elliptic functions.

The quantities λ and M in the relation (1) are defined by the formulas

$$\lambda = k^{2m+1}\prod_{r=1}^{m}C_{2r-1}^2,$$

$$M = \prod_{r=1}^{m}\frac{C_{2r-1}}{C_{2r}},$$

$$C_r = \operatorname{sn}^2\left(\frac{rK}{2m+1}\,;k\right)$$

where the complete elliptic integral of the first kind L corresponds to the modulus λ. Thus

$$L = \frac{K}{(2m+1)\,M}.$$

The formula (1) expresses the so-called division by $2m+1$ of the period (first period) of the function $\operatorname{dn}(u;k)$.

These formulas are analogous to the trigonometric equations

$$\cos n\varphi = \frac{1}{2}\{(x+\sqrt{x^2-1})^n + (x-\sqrt{x^2-1})^n\},$$

$$\cos\varphi = x,$$

which express the division by n of the period 2π of the function $\cos\varphi$.

Let us put

$$x = \operatorname{sn}^2(u;k)$$

and consider the function

$$y = \frac{2}{1+\sqrt{1-\lambda^2}}\operatorname{dn}\left(\frac{u}{M}\,;\lambda\right) = \frac{2}{1+\sqrt{1-\lambda^2}}\sqrt{1-k^2x}\prod_{r=1}^{m}\frac{1-k^2C_{2r-1}x}{1-k^2C_{2r}x}.$$

Suppose n increases from 0 to K, in other words, from 0 to $(2m+1)\,ML$. At the same time, x will vary from 0 to 1, but $\operatorname{dn}\left(\frac{u}{M}\,;\lambda\right)$ will always remain between 1 and $\sqrt{1-\lambda^2}$, so that y will lie between

$$\frac{2}{1+\sqrt{1-\lambda^2}} = 1 + \frac{1-\sqrt{1-\lambda^2}}{1+\sqrt{1-\lambda^2}} = 1+\mu$$

and

$$\frac{2\sqrt{1-\lambda^2}}{1+\sqrt{1-\lambda^2}} = 1 - \frac{1-\sqrt{1-\lambda^2}}{1+\sqrt{1-\lambda^2}} = 1-\mu \ .$$

Moreover, y will assume the value $1+\mu$ at the points $u = 0,\ 2ML,\ 4ML,\ \ldots$ $2mML$, and the value $1-\mu$ at the points $u = ML,\ 3ML,\ \ldots,\ (2m+1)ML$.

Hence the difference

$$1 - y(x)$$

attains its maximal value μ, in the interval $[0, 1]$, at the consecutive points

$$C_0 < C_1 < \ldots < C_{2m+1}$$

of this interval, where the sign of this difference is to be changed alternately at these points.

Since, in the case at hand, we have

$$N = 2m+2, \quad m = n, \ d = 0$$

and accordingly

$$N = m + n + 2 - d$$

it follows from P. L. Tchebysheff's general theorem that $y(x)$ is a function which, among all functions of the form

$$\sqrt{1-k^2x}\ \frac{q_0x^m + q_1x^{m-1} + \ldots + q_m}{p_0x^m + p_1x^{m-1} + \ldots + p_m}$$

deviates the least from unity, in the interval $[0, 1]$.

Consequently, the required relative deviation is equal to

$$\mu = \frac{1-\sqrt{1-\lambda^2}}{1+\sqrt{1-\lambda^2}},$$

with

$$\lambda = k^{2m+1}\prod_{r=1}^{m} \mathrm{sn}^4\left(\frac{2r-1}{2m+1}K; k\right).$$

It is to be noted that the problem considered in this section, and which was first solved by E. Solotareff and P. L. Tchebysheff [28], is of importance in the theory of electric filters (see the paper by Cauer [29]).

40. The Passage to Periodic Functions. Let us assume now that $f(\Theta)$ is a continuous periodic function with the period 2π. It is required to find a trigonometric sum of the order n, to wit

$$S_n(\Theta) = A_0 + A_1\cos\Theta + B_1\sin\Theta + \ldots + A_n\cos n\Theta + B_n\sin n\Theta$$

which yields the best approximation for the function $f(\Theta)$ along the entire number axis.

If we carry out the substitution of variables

$$x = \operatorname{tg} \frac{\Theta}{2}$$

the interval $-\pi \leq \Theta \leq \pi$ will be mapped upon the interval $-\infty \leq x \leq \infty$. Since we have

$$\cos \Theta = \frac{1 - x^2}{1 + x^2}, \quad \sin \Theta = \frac{2x}{1 + x^2}$$

and since $\cos k\,\Theta$ and $\dfrac{\sin (k+1)\,\Theta}{\sin \Theta}$ are polynomials of degree k in $\cos \Theta$ we obtain, after substitution,

$$S_n(\Theta) = \frac{p_0 x^{2n} + p_1 x^{2n-1} + \ldots + p_{2n}}{(1 + x^2)^n}.$$

Consequently, our problem leads to the problem of determining (in the interval $[-\infty, \infty]$) the best approximation of the function

$$F(x) = f(\Theta)$$

by means of expressions of the form

$$W_{2n}(x) = \frac{p_0 x^{2n} + p_1 x^{2n-1} + \ldots + p_{2n}}{(1 + x^2)^n}.$$

The expression $W_{2n}(x)$ may be regarded as a special case of the expression (T) in Section 31 for $m = 0$ and $s(x) = \dfrac{1}{(1 + x^2)^n}$. It is easy to see that the general theorem is applicable, and the theorem of P. L. Tchebysheff now reads: [3] *A trigonometric sum of order n which deviates the least from a given continuous periodic function throughout the number axis is uniquely and completely characterized by the fact that the number of consecutive points at which the difference* $f(\Theta) - S_n(\Theta)$ *with alternate change of sign assumes the value* $\max |f(\Theta) - S_n(\Theta)|$, *in the interval* $-\pi < \Theta \leq \pi$ *(or in any half-closed interval of length* 2π*), is not less than* $2n + 2$.

In addition, we will formulate the theorem of de la Vallée-Poussin: *If the difference* $f(\Theta) - S_n(\Theta)$ *assumes the values*

$$\lambda_1, -\lambda_2, \lambda_3, \ldots, -\lambda_{2n+2}$$

at $2n + 2$ *consecutive points of the interval* $-\pi < \Theta \leq \pi$ *(or of any half-closed interval of length* 2π*), where all the quantities* λ_k *are greater than zero, then the deviation of the best approximation of the function* $f(\Theta)$ *by means of trigonometric sums of order n is not less than* $\min \{\lambda_1, \lambda_2, \ldots, \lambda_{2n+2}\}$.

[3] This theorem can also be proved directly, and we recommend that the reader do so.

41. An Example of Approximating with the Aid of Periodic Functions. As a first example, we propose the problem of determining a trigonometric sum of the form

$$\alpha \cos n\Theta + \beta \sin n\Theta + a_{n-1} \cos (n-1)\Theta + b_{n-1} \sin (n-1)\Theta + \ldots + a_0$$

that deviates the least from zero, where α and β are given.

We are dealing here with the problem of finding the best approximation of a function

$$f(\Theta) = \alpha \cos n\Theta + \beta \sin n\Theta = \sqrt{\alpha^2 + \beta^2} \cos n(\Theta - \Theta_0)$$

by means of trigonometric sums of the $(n-1)$st order.

By virtue of the theorem in Section 40, the required trigonometric sum of the $(n-1)$st order is identically zero if the difference $f(\Theta) - 0$ assumes its maximal value $\sqrt{\alpha^2 + \beta^2}$, with alternate change of sign, at $2n$ consecutive points

$$\Theta_0 + \frac{k\pi}{n} \qquad (k = 1, 2, \ldots, 2n)$$

of the interval $\Theta_0 < \Theta \leq \Theta_0 + 2\pi$.

Hence the minimum deviation is equal to $\sqrt{\alpha^2 + \beta^2}$.

42. The Weierstrass Function. As a further example, let us consider the Weierstrass function

$$f(\Theta) = \sum_{k=0}^{\infty} a^k \cos (b^k \Theta),$$

where a is a positive number less than 1 and b is an odd integer greater than 1. Since the series which defines the function $f(\Theta)$ converges uniformly, it follows that $f(\Theta)$ is a continuous function whose period is obviously 2π.

For any natural number n, we can determine a natural number $k+1$ such that

$$b^k \leq n < b^{k+1}.$$

Let us consider the expression

$$S_n(\Theta) = \sum_{m=0}^{k} a^m \cos (b^m \Theta),$$

which represents a trigonometric sum of order $\leq n$.

The deviation of $S_n(\Theta)$ from $f(\Theta)$ is given by

$$\max |f(\Theta) - S_n(\Theta)| = \max \left| \sum_{m=k+1}^{\infty} a^m \cos (b^m \Theta) \right| = \sum_{m=k+1}^{\infty} a^m = \frac{a^{k+1}}{1-a} = L_n.$$

It is easy to see that the difference $f(\Theta) - S_n(\Theta)$ with sign that changes alternately at $2b^{k+1}$ consecutive points

$$\Theta_r = \frac{r\pi}{b^{k+1}} \qquad (r = 1, 2, \ldots, 2b^{k+1})$$

of the interval $0 < \Theta \leq 2\pi$, assumes the value L. But since $2b^{k+1} \geq 2n + 2$, it follows from the theorem in Section 40 that $S_n(\Theta)$ is a trigonometric sum which deviates less from $f(\Theta)$ than does any other trigonometric sum of the n-th order.

Thus we see that, for arbitrary values of n, the best approximation for $f(\Theta)$ is given by trigonometric sums which are the first n terms of the trigonometric series that defines the function $f(\Theta)$.

There are also other examples where this rare phenomenon occurs. In the example under consideration, we have

$$L_n = \frac{a^{k+1}}{1-a} < \frac{1}{1-a}\,\frac{1}{n^a}$$

where

$$a = \log \frac{1}{a} : \log b.$$

The foregoing follows from the fact that

$$a^{k+1} = b^{-a(k+1)} < \frac{1}{n^a}.$$

Later on, we shall have occasion to return to this valuation.

43. Haar's Problem. To P. L. Tchebysheff's domain of ideas, belongs an interesting general problem due to A. Haar [31], which we will now consider.

Suppose $f_1(P), \ldots, f_n(P)$ are n given linearly independent real functions of the point P of a bounded closed set $\mathfrak{M}$ in ordinary space of any number of dimensions. Let us consider the "polynomials"

$$F(P; x) = x_1 f_1(P) + x_2 f_2(P) + \ldots + x_n f_n(P) \ ;$$

we will call the quantity

$$\Delta(x) = L(x; f) = \max_{P \in \mathfrak{M}} |f(P) - F(P; x)|$$

the deviation of the polynomial $F(P; x)$ from a certain function $f(P)$ which is continuous on $\mathfrak{M}$.

Since the totality of all the functions which are continuous on $\mathfrak{M}$ is a linear normalized space (space C) in the light of our usual definition of distance, there exists, by reason of Section 8, a polynomial $F(P; x)$ corresponding to every function $f(P)$ such that $F(P; x)$ deviates the least from $f(P)$.

Haar's problem calls for the necessary and sufficient conditions which the functions $f_k(P)$ must satisfy in order that, for every function $f(P)$, the particular polynomial of least deviation be uniquely determined.

A. Haar's result is: *The polynomial of least deviation is uniquely determined for every function $f(P)$ if and only if every polynomial $F(P; x) \not\equiv 0$ in $\mathfrak{M}$ has no more than $n - 1$ distinct zeros.*

44. Proof of the Necessity of Haar's Condition. Let $F(P; x)$ be a polynomial that has $P_1, P_2, \ldots, P_n$ for its zeros in $\mathfrak{M}$.

Then we have

$$\begin{vmatrix} f_1(P_1) \ldots f_n(P_1) \\ \cdots\cdots\cdots \\ \cdots\cdots\cdots \\ f_1(P_n) \ldots f_n(P_n) \end{vmatrix} = 0 .$$

Hence there exists a non-trivial system of numbers $c_1, c_2, \ldots, c_n$ for which

$$c_1 f_k(P_1) + c_2 f_k(P_2) + \ldots + c_n f_k(P_n) = 0 \qquad (k = 1, 2, \ldots, n) .$$

Whence it follows that the equation

$$c_1 F(P_1; x) + c_2 F(P_2; x) + \ldots + c_n F(P_n; x) = 0 \tag{1}$$

is valid for every polynomial $F(P; x)$.

Let us determine a number λ such that

$$\max_{P \in \mathfrak{M}} | \lambda F(P; \alpha) | < 1 .$$

Let $g(P)$ be any function which is continuous in $\mathfrak{M}$ and such that

$$| g(P) | \leqq 1 \qquad (P \in \mathfrak{M})$$

and

$$g(P_i) = \operatorname{sign} c_i, \qquad \text{for } c_i \neq 0 \ (i = 1, 2, \ldots, n) .$$

The function

$$f(P) = g(P) \{1 - | \lambda F(P; \alpha) |\}$$

also possesses this property.

We are going to show that there is an infinite number of polynomials having the least deviation from the function $f(P)$. In fact, the quantity $L(x; f)$, for arbitrary values $x_1, x_2, \ldots, x_n$, is not less that 1, since otherwise, by virtue of the equations

$$f(P_i) = \operatorname{sign} c_i, \qquad c_i \neq 0 \quad (i = 1, 2, \ldots, n) ,$$

the corresponding equations

$$\operatorname{sign} F(P_i; x) = \operatorname{sign} f(P_i) = \operatorname{sign} c_i$$

would hold for all the quantities c_i which differ from zero, in contradiction of the relation (1).

On the other hand, when $|\varepsilon| \leqq 1$, we have

$$\begin{aligned} | f(P) - \varepsilon \lambda F(P; \alpha) | &\leqq | f(P) | + | \varepsilon \lambda F(P; \alpha) | = \\ &= | g(P) | \{1 - | \lambda F(P; \alpha) |\} + | \varepsilon \lambda F(P; \alpha) | \leqq \\ &\leqq 1 - | \lambda F(P; \alpha) | + | \varepsilon \lambda F(P; \alpha) | = \\ &= 1 - (1 - | \varepsilon |) | \lambda F(P; \alpha) | \leqq 1 , \end{aligned}$$

so that, corresponding to every $\varepsilon (-1 \leqq \varepsilon \leqq 1)$, $\varepsilon \lambda F(P; \alpha)$ is a polynomial of least deviation from $f(P)$ (since the deviation is equal to unity).

The necessity of Haar's condition is thus proven.

45. Proof of the Sufficiency of Haar's Condition. Let us assume that Haar's condition is fulfilled and let us prove three lemmas.

Lemma 1. *If*

$$(1) \qquad \begin{vmatrix} f_i(P_i) & f_{i+1}(P_i) & \dots & f_k(P_i) \\ \dots & \dots & \dots & \dots \\ f_i(P_k) & f_{i+1}(P_k) & \dots & f_k P(k) \end{vmatrix} \neq 0 \quad (1 \leqq i < k < n),$$

then, corresponding to any value $q (k < q \leq n)$, *we can find such points*

$$P_{k+1}, P_{k+2}, \dots, P_q$$

that we have

$$\begin{vmatrix} f_i(P_i) & f_{i+1}(P_i) & \dots & f_q(P_i) \\ \dots & \dots & \dots & \dots \\ f_i(P_q) & f_{i+1}(P_q) & \dots & f_q(P_q) \end{vmatrix} \neq 0'.$$

Proof: By virtue of (1), the non-trivial polynomial

$$\Phi(P) = \begin{vmatrix} f_i(P_i) & \dots & f_k(P_i) & f_{k+1}(P_i) \\ \dots & \dots & \dots & \dots \\ f_i(P_k) & \dots & f_k(P_k) & f_{k+1}(P_k) \\ f_i(P) & \dots & f_k(P) & f_{k+1}(P) \end{vmatrix}$$

cannot have more than $n - 1$ zeros. Hence there exists a point P_{k+1} for which

$$\Phi(P_{k+1}) \neq 0|.$$

By repeating this reasoning, we will prove the following lemma:

Lemma 2. *If the points* $P_1, P_2, \dots, P_k$ $(k < n)$ *are all distinct, then at least one of the determinants of the matrix of k-th degree*

$$\left\| \begin{matrix} f_1(P_1) & \dots & f_n(P_1) \\ \dots & \dots & \dots \\ f_1(P_k) & \dots & f_n(P_k) \end{matrix} \right\|$$

is different from zero.

Proof: Let us first take $k = 1$. We will show that at least one of the quantities $f_i(P_1)$, $(i = 1, 2, \dots, n)$ is different from zero. For, if such is not the case, let us determine points $Q_2, \dots, Q_n$ so as to have

$$\begin{vmatrix} f_2(Q_2) & \dots & f_n(Q_2) \\ \dots & \dots & \dots \\ f_2(Q_n) & \dots & f_n(Q_n) \end{vmatrix} \neq 0$$

(to that end, it suffices to take any point Q_2, provided that $f(Q_2) \neq 0$ and Lemma 1 is fulfilled); we thus obtain a non-trivial polynomial

$$\begin{matrix} f_1(P) & f_2(P) & \cdots & f_n(P) \\ f_1(Q_2) & f_2(Q_2) & \cdots & f_n(Q_2) \\ \cdots & \cdots & \cdots & \cdots \\ f_1(Q_n) & f_2(Q_n) & \cdots & f_n(Q_n) \end{matrix}$$

with n distinct zeros at $P_1, Q_2, \ldots, Q_n$, respectively.

The case $k > 1$ will be established by the method of induction. Suppose the lemma is true for $k = 1, 2, \ldots, m-1$. Without any loss of generality, we may assume that

$$\begin{vmatrix} f_2(P_2) & \cdots & f_m(P_2) \\ \cdots & \cdots & \cdots \\ f_2(P_m) & \cdots & f_m(P_m) \end{vmatrix} \neq 0 .$$

By virtue of Lemma 1, we can find points $Q_{m+1}, \ldots, Q_n$ such that

$$\begin{vmatrix} f_2(P_2) & \cdots & f_n(P_2) \\ \cdots & \cdots & \cdots \\ f_2 Q_n) & \cdots & f_n(Q_n) \end{vmatrix} \neq 0.$$

Then, however,

$$\begin{vmatrix} f_1(P) & \cdots & f_n(P) \\ f_1(P_2) & \cdots & f_n(P_2) \\ \cdots & \cdots & \cdots \\ f_1(Q_n) & \cdots & f_n(Q_n) \end{vmatrix}$$

is a non-trivial polynomial. If our proposition were false, this polynomial would have a zero at each of the n points $P_1, P_2, \ldots, P_m, Q_{m+1}, \ldots, Q_n$, which is impossible.

Lemma 3. *If the number of points of the set* $\mathfrak{M}$ *where*

$$|f(P) - F(P; x)| = L(x) \equiv L(x; f) \tag{2}$$

is less than n, *then* $F(P; x)$ *is not a polynomial that deviates least from* $f(P)$.

Proof: Let $P_1, \ldots, P_m (m < n)$ be wholly distinct points of $\mathfrak{M}$ for which (2) is valid. According to the second of these lemmas, the system of equations

$$f_1(P_k)\,\xi_1 + f_2(P_k)\,\xi_2 + \ldots + f_n(P_k)\,\xi_n = f(P_k) - F(P_k; x)$$
$$(k = 1, 2, \ldots, m)$$

can be solved for $\xi_1, \ldots, \xi_n$.

Let

$$R(P) = f(P) - F(P; x).$$

We now choose in the vicinity of each of the points P_k $(k = 1, 2, \ldots, m)$ a region $\mathfrak{M}_k$ small enough so that

$$\mu_k = \min_{P \in \mathfrak{M}_k} |R(P)| > 0, \quad \min_{P \in \mathfrak{M}_k} |F(P; \xi)| \geqq \frac{L(x)}{2} .$$

Furthermore, let

$$M_k = \max_{P \in \mathfrak{M}_k} |F(P; \xi)|,$$

$$M = \max_{P \in \mathfrak{M}^*} |F(P; \xi)|, \quad L^*(x) = \max_{P \in \mathfrak{M}^*} |R(P)|,$$

with $\mathfrak{M}^* = \mathfrak{M} - \mathfrak{M}_1 - \mathfrak{M}_2 - \ldots - \mathfrak{M}_m$.

Evidently,

$$\mu = L(x) - L^*(x) > 0.$$

Let us take a number ε which satisfies the inequality

$$0 < \varepsilon < \min\left\{\frac{\mu}{M}, \frac{\mu_1}{M_1}, \ldots, \frac{\mu_m}{M_m}\right\}$$

and put

$$x_i' = x_i + \varepsilon\xi_i \qquad (i = 1, 2, \ldots, n).$$

We then have

$$|f(P) - F(P; x')| = |f(P) - F(P; x) - \varepsilon F(P; \xi)| = \\ = |R(P) - \varepsilon F(P; \xi)|,$$

whence we obtain the inequalities

$$|f(P) - F(P; x')| \leqq |R(P)|\left\{1 - \varepsilon\frac{F(P; \xi)}{R(P)}\right\} \leqq L(x)\left(1 - \frac{\varepsilon}{2}\right)$$

for

$$P \in \mathfrak{M}_k \qquad (k = 1, 2, \ldots, m)$$

and

$$|f(P) - F(P; x')| \leqq |R(P)| + \varepsilon |F(P; \xi)| \leqq L^*(x) + \varepsilon M < L(x)$$

for

$$P \in \mathfrak{M}^*.$$

Accordingly,

$$L(x') = \max_{P \in \mathfrak{M}} |f(P) - F(P; x')| < L(x),$$

which completes the proof of Lemma 3.

It is now easy to demonstrate the sufficiency of Haar's condition. Let us assume that there are two distinct polynomials

$$F(P; x), \; F(P; y),$$

each of which yields the minimum deviation for the function $f(P)$.

Then the polynomial $F\left(P; \frac{x+y}{2}\right)$ would be endowed with the same property, since

$$\left|F\left(P; \frac{x+y}{2}\right) - f(P)\right| \leqq \frac{1}{2}|F(P; x) - f(P)| + \frac{1}{2}|F(P; y) - f(P)|.$$

By reason of Lemma 3, the equation

$$\left| f(P) - F\left(P; \frac{x+y}{2}\right) \right| = L$$

with $L = L\left(\frac{x+y}{2}\right) = L(x) = L(y)$ has at least n distinct zeros $P_1, P_2, \ldots, P_n$ in $\mathfrak{M}$. But, in order that the equation

$$\left| f(P_i) - F\left(P_i; \frac{x+y}{2}\right) \right| = L$$

be valid, it is necessary that the equations

$$f(P_i) - F(P_i; x) = f(P_i) - F(P_i; y) = \pm L$$

be valid. Hence it follows from our assumption that the non-trivial polynomial

$$F(P; x - y)$$

has n distinct zeros in $\mathfrak{M}$. Since this is impossible, we have proved the sufficiency of Haar's condition.

46. An Example Related to Haar's Problem [32]. Let us consider an infinite (either denumerable or non-denumerable) system of linear equations in n unknowns:

$$(1) \qquad \begin{cases} a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = l, \\ a_1' x_1 + a_2' x_2 + \ldots + a_n' x_n = l', \\ a_1'' x_1 + a_1'' x_2 + \ldots + a_n'' x_n = l'', \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{cases}$$

where the coefficients and the quantities on the right-hand side are bounded and real, and let us assume further that these equations are inconsistent.

It would be senseless to ask for an exact solution of the system (1); however, one might properly inquire about approximate solutions of these equations. As an example, let us try to find quantities $x_1, x_2, \ldots, x_n$ giving the minimum deviation

$$L(x) = \sup_{(a)} \left| l^{(a)} - \sum_{k=1}^{n} a_k^{(a)} x_k \right| .$$

This problem fits in with Haar's general problem. Let us take the above mentioned numbers

$$a_1^{(a)}, a_2^{(a)}, \ldots, a_n^{(a)}, l^{(a)}$$

as coordinates of a certain point P. They are continuous functions of P. Hence the system (1) can be written also in the form

$$x_1 f_1(P) + x_2 f_2(P) + \ldots + x_n f_n(P) = f(P)$$

If the point P, corresponding to the equations (1), belongs to the set $\mathfrak{M}$, which by hypothesis is bounded, then we are concerned with the determination of the quantities x_i, subject to the condition that the quantity

$$L(x) = \sup_{P \in \mathfrak{M}} \left| f(P) - \sum_{k=1}^{n} x_k f_k(P) \right| \tag{2}$$

has the least possible value.

Obviously $L(x)$ will remain unaltered if we close the set $\mathfrak{M}$. Hence we may assume from the start that $\mathfrak{M}$ is a closed bounded set, and hence we may replace (2) by

$$L(x) = \max_{P \in \mathfrak{M}} \left| f(P) - \sum_{k=1}^{n} x_k f_k(P) \right| .$$

We know that there exists a system of values x_i for which the quantity $L(x)$ assumes the smallest value. By virtue of Haar's theorem, this system will be uniquely determined for any system of numbers $l^{(a)}$ if, corresponding to arbitrarily chosen $\xi_i \left(\sum_{k=1}^{n} \xi_k^2 > 0 \right)$, the polynomial

$$\xi_1 f_1(P) + \xi_2 f_2(P) + \ldots + \xi_n f_n(P) \cdot$$

has no more than $n - 1$ distinct roots in $\mathfrak{M}$, that is to say if there does not exist a non-trivial system $\xi_1, \xi_2, \ldots, \xi_n$ for which the linear forms

$$\sum_{k=1}^{n} a_i^{(a)} \xi_i$$

have more than $n - 1$ zeros; in other words, if none of the determinants

$$\begin{vmatrix} a_1^{(i_1)} & a_2^{(i_1)} & \ldots & a_n^{(i_1)} \\ a_1^{(i_2)} & a_2^{(i_2)} & \ldots & a_n^{(i_2)} \\ \cdot\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot\cdot \\ a_1^{(i_n)} & a_2^{(i_n)} & \ldots & a_n^{(i_n)} \end{vmatrix}$$

is equal to zero.

47. P. L. Tchebysheff's Systems of Functions. Let us consider a system of real functions

$$f_1(x), f_2(x), \ldots, f_n(x) \tag{S}$$

which are continuous in the finite [4] interval $[a, b]$ and which satisfy Haar's condition. Following S. N. Bernstein [33], we call this system briefly *Tchebysheff's system with respect to the interval* $[a, b]$.

[4] The assumption that the interval $[a, b]$ be finite may be omitted.

Lemma. If $x_1, x_2, \ldots, x_{n-1}$ are any distinct points of the interval $[a, b]$, there exists one (and, except for a constant factor, only one) non-trivial polynomial

$$F(x; \lambda) = \lambda_1 f_1(x) + \lambda_2 f_2(x) + \ldots + \lambda_n f_n(x),$$

which has $x_1, x_2, \ldots, x_{n-1}$ for its zeros.

If x_k lies inside the interval (a, b), the polynomial changes sign on passing through x_k.

Proof: The polynomial

$$D(x; x_1, x_2, \ldots, x_{n-1}) = \begin{vmatrix} f_1(x) & f_2(x) & \ldots f_n(x) \\ f_1(x_1) & f_2(x_1) & \ldots f_n(x_1) \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_1(x_{n-1}) & f_2(x_{n-1}) & \ldots f_n(x_{n-1}) \end{vmatrix},$$

which is non-trivial (cf. Lemma 2, Section 45), vanishes at the points $x_1, x_2, \ldots, x_{n-1}$ and has no other zeros, since (S) is a Tchebysheff system.

Let us construct a system of linear equations which determines the coefficients of the polynomial $F(x; \lambda)$; then we see that

$$F(x; \lambda) \equiv C D(x; x_1, x_2, \ldots, x_{n-1})$$

is true, where C is a constant.

Since the inequality

$$(1) \qquad (a \leqq) x_0 < x_1 < \ldots < x_{n-1} (\leqq b)$$

is fulfilled whenever

$$D(x_0; x_1, x_2, \ldots, x_{n-1}) \neq 0$$

it follows that, by virtue of (1), the determinant

$$D(x_0; x_1, x_2, \ldots, x_{n-1})$$

always has one and the same sign.

Whence it follows that for the points x_k which lie within the interval (a, b) the numbers

$$D(x_k - \varepsilon; x_1, x_2, \ldots, x_{n-1}), \quad D(x_k + \varepsilon; x_1, x_2, \ldots, x_{n-1})$$
$$(1 \leqq k \leqq n - 1)$$

always have opposite signs for sufficiently small values of $\varepsilon > 0$. Thus our lemma is established.

48. Generalization of P. L. Tchebysheff's Theorem. In this section our objective will be the following generalization of Tchebyheff's theorem of Section 35. *If* (S), Section 47, *is a Tchebysheff system with respect to the interval $[a, b]$, and $f(x)$ is any function continuous in $[a, b]$, then the polynomial $F(x; \alpha)$ which deviates the least from $f(x)$ in $[a, b]$ in the metric space C (we have already proven the existence and uniqueness of this polynomial* [5]*), is completely determined by the*

fact that the difference $f(x) - F(x; \alpha)$ *attains its maximal value at no less than* $n + 1$ *consecutive points of the interval* $[a, b]$, *the sign of this difference at those points being taken alternately plus and minus.*

Let us assume that the difference

$$f(x) - F(x; \alpha)$$

takes on the value

$$L = \max_{a \leqq x \leqq b} |f(x) - F(x; \alpha)|$$

with alternate change in sign at $q \leqq n$ consecutive points $y_1 < y_2 < \ldots < y_q$ of the interval $[a, b]$.

Obviously, the interval $[a, b]$ can always be divided into q subintervals

$$[a, x_1], [x_1, x_2], \ldots, [x_{q-1}, b] \tag{1}$$

where the quantities involved in the relation

$$a \leqq y_1 < x_1 < \ldots < y_{q-1} < x_{q-1} < y_q \leqq b$$

are so chosen that, in turn, one of the two inequalities

$$-L \leqq f(x) - F(x; \alpha) < L - \mu, \quad -L + \mu < f(x) - F(x; \alpha) \leqq L$$

subject to the conditions $0 < \mu < \frac{L}{2}$ and

$$f(x_{q-1}) = F(x_{q-1}; \alpha)$$

is satisfied in the intervals (1). Let us chose a point X between the points x_{q-1} and y_q so that in the interval

$$[x_{q-1}, X] \tag{2}$$

the inequality

$$-L + \mu < f(x) - F(x; \alpha) < L - \mu$$

is true, and let us choose in the interval (x_{q-1}, X) any points

$$x_q < x_{q+1} < \ldots < x_{q+2m-1}$$

with $m = \left[\frac{n-q}{2}\right]$ and hence $q + 2m - 1$ equal to either $n - 1$ or $n - 2$.

We now construct the polynomial

$$F(x; \beta) = F(x; \alpha) + \varepsilon D(x; x_1, x_2, \ldots, x_{n-1})$$

with $x_{n-1} = b$, for $q + 2m - 1 = n - 2$.

Let us choose a number ε so that

$$\max_{a \leqq x \leqq b} | \varepsilon D(x; x_1, x_2, \ldots, x_{n-1}) | < \mu .$$

Obviously, ε may be chosen such that the condition

---

[5] We recommend that at this time the reader carry out the proof for the instant case so as to be convinced that the assumption of the finiteness of the interval $[a, b]$ is not essential.

$$\text{sign } \{f(y_1) - F(y_1; \alpha)\} = \text{sign } \varepsilon D\,(x; x_1, x_2, \ldots, x_{n-1})$$

is fulfilled in the interval $[a, x_1]$. Then, by virtue of the lemma in the previous section, we have

$$\text{sign } \varepsilon D\,(x; x_1, x_2, \ldots, x_{n-1}) = \text{sign } \{f(y_k) - F(y_k; \alpha)\}$$

subject to the conditions

$$x_{k-1} \leqq x \leqq x_k \text{ for } k = 1, 2, \ldots, q-1 \text{ and } x_0 = a.$$

Whence it follows that the inequality

$$|f(x) - F(x; \beta)| < L$$

holds in the interval $[a, x_{q-1}]$.

On the other hand, this inequality holds in the interval $[x_{q-1}, b]$ if $q + 2m - 1 = n - 1$. Consequently we have shown for this case that $F(x; \alpha)$ is not the polynomial which deviates the least from $f(x)$.

If $q + 2m - 1 = n - 2$, then, in all cases where $a \leqq x < b$, we have

$$|f(x) - F(x; \beta)| < L$$

while

$$|f(b) - F(b; \beta)| \leqq L$$

If the sign in the last case is $<$, then $F(x; \beta)$ yields a better approximation than $F(x; \alpha)$.

If

$$|f(b) - F(b; \beta)| = L,$$

it suffices to take any polynomial $F(x; \gamma)$ for which [6]

$$F(b; \gamma)\,[f(b) - F(b; \beta)] > 0$$

in order that, for a certain $\delta > 0$, the inequality

$$|f(x) - F(x; \beta) - \delta F(x; \gamma)| < L$$

shall hold for every interval $a \leqq x \leqq b$.

Thus the necessity of the criterion stated in the theorem has been proved. The proof of the sufficiency of this criterion is easy and is left to the reader. In addition, we leave it to the reader to generalize the theorem of de la Vallée-Poussin, along the lines indicated here [33].

49. On a Question Pertaining to the Approximation of a Continuous Function in the Space L. Suppose

$$f(x), f_1(x), f_2(x), \ldots, f_n(x)$$

are given, real-valued functions continuous in the finite interval $[a, b]$. We know that there exists a polynomial

$$F(x; \lambda) = \lambda_1 f_1(x) + \lambda_2 f_2(x) + \ldots + \lambda_n f_n(x)$$

[6] It must be borne in mind that not all the functions $f_k(x)$ $(k = 1, 2, \ldots, n)$ can be zero for $x = b$, since we have a Tchebysheff system with respect to the closed interval $[a, b]$.

for which the integral

$$I[f;\alpha]=\int_a^b |f(x)-F(x;\alpha)|\,dx$$

assumes its least value $I[f]$.

If the functions

$$f_1(x), f_2(x), \ldots, f_n(x) \tag{1}$$

constitute a Tchebysheff system with respect to the interval $[a, b]$, we will show that this polynomial furnishing the best approximation is uniquely determined for every function $f(x)$. This theorem was proved first by D. Jackson [34], but special cases thereof were already familiar to Tchebysheff.

First, we will prove two lemmas.

Lemma 1. *If a certain continuous function $f(x)$ has two polynomials*

$$F(x;\lambda),\quad F(x;\mu)$$

which furnish the best approximation (in the metric of L), then the inequality

$$\{f(x)-F(x;\lambda)\}\left\{f(x)-F\left(x;\frac{\lambda+\mu t}{1+t}\right)\right\}\geqq 0,$$

where $t\geqq 0$, is valid at every point of the interval $[a, b]$.

Proof: Obviously, the relation

$$I\left[f;\frac{\lambda+\mu t}{1+t}\right]\leqq\frac{1}{1+t}I[f;\lambda]+\frac{t}{1+t}I[f;\mu]=I[f].$$

is true for any number $t\geqq 0$.

Now, from the definition of the quantity $I[f]$, it follows that

$$I\left[f;\frac{\lambda+\mu t}{1+t}\right]\geqq I[f]\ .$$

Hence, for any value of $t\geqq 0$,

we have

$$I\left[f;\frac{\lambda+\mu t}{1+t}\right]=I[f]\ .$$

Consequently, for every value of $t\geqq 0$, the polynomial $F\left(x;\frac{\lambda+\mu t}{1+t}\right)$, as well as the polynomials $F(x;\lambda)$ and $F(x;\mu)$, yields the best approximation of the function $f(x)$.

Hence, for $\sigma=\frac{\lambda+\mu t}{1+t}$, we have

$$\int_a^b\left|f(x)-F\left(x;\frac{\lambda+\sigma}{2}\right)\right|dx=$$

$$=\frac{1}{2}\int_a^b|f(x)-F(x;\lambda)|\,dx+\frac{1}{2}\int_a^b|f(x)-F(x;\sigma)|\,dx.$$

Therefore the differences

$$R(x;\lambda) = f(x) - F(x;\lambda), \quad R(x;\sigma) = f(x) - F(x;\sigma)$$

have one and the same sign at every point of the interval $[a, b]$.

Lemma 2. If $F(x;\lambda)$ is a polynomial furnishing the best approximation for $f(x)$ and the difference $R(x;\lambda)$ changes sign alternately at $p \geq n$ points of the interval (a, b), then $F(x;\lambda)$ is the only polynomial furnishing the best approximation for $f(x)$.

Proof: Suppose the lemma is untrue, and let the aforementioned points be denoted by

$$x_1 < x_2 < \ldots < x_p \qquad (a < x_1, x_p < b). \tag{2}$$

In accordance with Lemma 1, the expression $R(x;\mu)$, which has already been set up for the second polynomial furnishing the best approximation, will also change sign alternately at the points (2).

Consequently, we have

$$F(x;\lambda) = F(x;\mu),$$

at those points; that is, the non-trivial polynomial

$$F(x;\lambda - \mu)$$

has $p \geq n$ zeros (2) in $[a, b]$. But that is impossible, because (1) is a Tchebysheff system.

We now come to the proof of D. Jackson's theorem.

Suppose $F(x;\lambda)$ and $F(x;\mu)$ are two distinct polynomials, each of which furnishes the best approximation of $f(x)$, and suppose, in conformity with Lemmas 1 and 2, each of the differences $R(x;\lambda)$ and $R(x;\mu)$ changes sign alternately at $q \leq n - 1$ points of the interval (a, b), in an identical manner.

Let these points be denoted by

$$x_1 < x_2 < \ldots < x_q \qquad (a < x_1, x_q < b)$$

and let the interval

$$[b - \delta, b] \tag{3}$$

be chosen so small that

$$b - \delta > x_q \ .$$

Let us choose any points

$$x_{q+1} < x_{q+2} < \ldots < x_{n-1}$$

in the interval (3), and let us construct the polynomial

$$F(x;\alpha) = C\,D(x; x_1, x_2, \ldots, x_{n-1}),$$

where C is so chosen that

$$\max_{a \leqq x \leqq b} |F(x; \alpha)| = 1,$$

$$\operatorname*{sign}_{a<x<a_1} F(x; \alpha) = \operatorname*{sign}_{a<x<x_1} R(x; \lambda) .$$

Then we shall have $[a, x_{q-1}]$

$$\operatorname{sign} F(x; \alpha) = \operatorname{sign} R(x; \lambda) = \operatorname{sign} R(x; \mu)$$

everywhere in $[a, x_{q-1}]$.

Let r_ε, s_ε, w_ε, where ε is any number > 0, denote the sets of those points of the interval $[a, b]$ at which the corresponding inequalities

$$(r_\varepsilon) \qquad \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) \right| \geqq \varepsilon, \quad R\left(x; \frac{\lambda + \mu t}{1+t}\right) F(x; \alpha) \geqq 0,$$

$$(s_\varepsilon) \qquad \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) \right| < \varepsilon, \quad R\left(x; \frac{\lambda + \mu t}{1+t}\right) F(x; \alpha) \geqq 0,$$

$$(w_\varepsilon) \qquad R\left(x; \frac{\lambda + \mu t}{1+t}\right) F(x; \alpha) < 0$$

are satisfied.

It is evident that the sum of these sets fills the entire interval and that the measure of the set w_ε does not exceed δ since w_ε lies in the interval (3).

Now, we have the relations

$$\int_{r_\varepsilon} \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) - \varepsilon F(x; \alpha) \right| dx =$$

$$= \int_{r_\varepsilon} \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) \right| dx - \varepsilon \int_{r_\varepsilon} |F(x; \alpha)| \, dx,$$

$$\int_{s_\varepsilon + w_\varepsilon} \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) - \varepsilon F(x; \alpha) \right| dx \leqq$$

$$\leqq \int_{s_\varepsilon + w_\varepsilon} \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) \right| dx + \varepsilon \int_{s_\varepsilon + w_\varepsilon} |F(x; \alpha)| \, dx.$$

Whence it follows that

$$\int_a^b \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) - \varepsilon F(x; \alpha) \right| dx \leqq$$

$$\leqq \int_a^b \left| R\left(x; \frac{\lambda + \mu t}{1+t}\right) \right| dx - \varepsilon \int_{r_\varepsilon} |F(x; \alpha)| \, dx + \varepsilon \int_{s_\varepsilon + w_\varepsilon} |F(x; \alpha)| \, dx.$$

Evidently,

$$\int_a^b \left| R\left(x; \frac{\lambda+\mu t}{1+t}\right) - \varepsilon F(x;\alpha) \right| dx \geqq \int_a^b R\left(x; \frac{\lambda+\mu t}{1+t}\right) dx.$$

Hence

$$\int_{s_\varepsilon+w_\varepsilon} |F(x;\alpha)|\, dx \geqq \int_{r_\varepsilon} |F(x;\alpha)|\, dx,$$

and consequently we have

$$\int_a^b |Fx;\alpha)|\, dx \leqq 2 \int_{s_\varepsilon+w_\varepsilon} |F(x;\alpha)|\, dx. \tag{4}$$

In consequence of the general reasoning in Section 8, there always exists a constant M such that the relation

$$\int_a^b |F(x;\beta)|\, dx = 1$$

implies the inequality [7]

$$|\beta_i| \leqq M \qquad (i = 1, 2, \ldots, n)$$

[7] In fact, the expression

$$\int_a^b |F(x,\gamma)|\, dx$$

has a certain minimum $m > 0$ on the sphere

$$|\gamma_1|^2 + |\gamma_2|^2 + \ldots + |\gamma_n|^2 = 1 .$$

Consequently, the inequality

$$\int_a^b |F(x;\beta)|\, dx \geqq m \sqrt{|\beta_1|^2 + |\beta_2|^2 + \ldots + |\beta_n|^2}.$$

is valid for arbitrary values of $\beta_1, \beta_2, \ldots, \beta_n$.
Hence, from the equation

$$\int_a^b |F(x;\beta)|\, dx = 1,$$

it follows that

$$|\beta_i| \leqq \frac{1}{m} = M .$$

and the further inequality

$$\max_{a \leqq x \leqq b} | F(x; \beta) | \leqq n M N$$

where N is the largest of the maxima of the absolute values of $f_k(x)$.

Hence the inequality (4) gives

$$G \leqq 2nMNG\,(\delta + \text{meas. } s_\varepsilon)$$

or

$$\frac{1}{2nMN} \leqq \delta + \text{meas. } s_\varepsilon \tag{5}$$

(if we take

$$G = \int_a^b | F(x; \alpha) | \, dx$$

so that $| \alpha_i | \leq MG$ and consequently $\max_{a \leqq x \leqq b} | F(x; \alpha) | \leq nMNG$).

Since the quantities M, N do not depend on the points $x_{q+1}, \ldots, x_{n-1}$ but only G does, and G is not present in (5), we can choose δ so small that the inequality

$$\frac{1}{2nMN} - \delta = h > 0$$

is satisfied. Consequently we have:

$$\text{meas. } s_\varepsilon \geq h.$$

If ε tends towards zero, we obtain

$$\text{meas. } s \geq h,$$

where s is the set of points of the interval $[a, b]$ at which

$$R\left(x; \frac{\lambda + \mu t}{1 + t}\right) = 0 .$$

Since the measure of the set s is not greater than h for any value of $t > 0$, there exist two distinct values of t such that the corresponding sets have in common a subset of positive measure; in other words, for certain values $t' > 0$, $t'' > 0$ $(t' \neq t'')$, the equation

$$F\left(x; \frac{\lambda + \mu t'}{1 + t'}\right) = F\left(x; \frac{\lambda + \mu t''}{1 + t''}\right)$$

holds in a set of positive measure. But that is absurd, since the polynomial

$$F\left(x; \frac{\lambda + \mu t''}{1 + t''} - \frac{\lambda + \mu t'}{1 + t'}\right)$$

cannot have more than $n - 1$ zeros in (a, b).

50. A. A. Markoff's Theorem. Let

$$(\Sigma) \qquad f_1(x), f_2(x), \ldots, f_n(x) \qquad (n \leqq \infty)$$

be given real-valued functions in the finite interval $[a, b]$. We say that Σ is a Markoff system of functions with respect to the interval $[a, b]$ if, for any (finite) $k \leq n$, the functions

$$f_1(x), f_2(x), \ldots, f_k(x)$$

form a Tchebysheff system with respect to the interval $[a, b]$.

As we know, corresponding to any $k < n$, there exists in the metric space L just one polynomial

$$F_k(x) = a_{k1}f_1(x) + a_{k2}f_2(x) + \ldots + a_{kk}f_k(x)$$
$$(k = 1, 2, \ldots)$$

with the least deviation from $f_{k+1}(x)$. Let us set

$$Q_1(x) = f_1(x),$$
$$Q_k(x) = f_k(x) - F_{k-1}(x) \qquad (k = 2, 3, \ldots).$$

It is easy to show that the polynomial $Q_k(x)$ changes sign alternately exactly $k - 1$ times inside the interval (a, b). As to the polynomial $Q_1(x)$, that is obvious. Let us consider $Q_k(x)$ when $k > 1$. Suppose $Q_k(x)$ changes sign alternately at j points

$$x_1 < x_2 < \ldots < x_j$$

of the interval (a, b). Because of the nature of a Tchebysheff system, we have $j \leq k - 1$. For the sake of definiteness, we may assume that

$$I_k = \int_a^b |Q_k(x)|\,dx = \int_a^{x_1} Q_k(x)\,dx - \int_{x_1}^{x_2} Q_k(x)\,dx + \ldots + (-1)^j \int_{x_j}^b Q_k(x)\,dx .$$

By virtue of the extremal property of the polynomial $F_{k-1}(x)$, the equations

$$\frac{\partial I_k}{\partial a_{k-1,r}} = 0 \qquad (r = 1, 2, \ldots, k - 1).$$

are valid. By differentiation, we obtain

$$\int_a^{x_1} f_r(x)\,dx - \int_{x_1}^{x_2} f_r(x)\,dx + \ldots + (-1)^j \int_{x_j}^b f_r(x)\,dx = 0$$
$$(r = 1, 2, \ldots, k - 1).$$

This relation can be written in the form

$$(1) \qquad \int_a^b f_r(x) \operatorname{sign} Q_k(x)\,dx = 0 \qquad (r = 1, 2, \ldots, k - 1).$$

Suppose, now, that $j \leq k-2$. Then, by reason of the lemma in Section 47, there exists a polynomial

$$G(x) = \lambda_1 f_1(x) + \lambda_2 f_2(x) + \ldots + \lambda_{j+1} f_{j+1}(x),$$

which changes sign alternately at the points

$$x_1, x_2, \ldots, x_j$$

and at no others. Therefore we should have

$$\int_a^b G(x) \operatorname{sign} Q_k(x)\, dx \neq 0,$$

which is impossible because of (1).

Let us now construct polynomials

$$P_k(x) = \beta_{k1} f_1(x) + \beta_{k2} f_2(x) + \ldots + \beta_{kk} f_k(x) \quad (k = 1, 2, 3, \ldots),$$

which satisfy the following relations of orthogonality:

$$(2) \qquad \begin{cases} \displaystyle\int_a^b P_k(x) \operatorname{sign} Q_r(x)\, dx = 0 & (r = 1, 2, \ldots, k-1), \\ \displaystyle\int_a^b P_k(x) \operatorname{sign} Q_k(x)\, dx = 1. \end{cases}$$

It is not difficult to see that the polynomials $P_k(x)$ are uniquely defined by the conditions (2).

The orthogonality relations (1) and (2) can be combined:

$$(3) \qquad \int_a^b P_k(x) \operatorname{sign} Q_r(x)\, dx = \begin{cases} 0 & (k \neq r), \\ 1 & (k = r) \end{cases} \quad (k, r = 1, 2, \ldots).$$

In conformity with the customary terminology, we can say that the sequence of polynomials $P_k(x)$ and the functions sign $Q_k(x)$ form a biorthogonal system. The biorthogonal system thus constructed enables us to associate with every (integrable) function $g(x)$ a series of polynomials $P_k(x)$:

$$g(x) \sim \sum_1^n A_k P_k(x),$$

where

$$A_k = \int_a^b g(x) \operatorname{sign} Q_k(x)\, dx.$$

The initial segments of this series can be employed to represent approximately the function $g(x)$.

P. L. Tchebysheff was the first to study similar developments [35]. Without dwelling on this matter any further, (cf. "Addenda and Problems," Sections 14, 15, 16), we pass to the theorem of A. A. Markoff [36]. This says: *Suppose that $f(x)$ is a function continuous in the interval $[a, b]$, and that the coefficients of the polynomial*

$$F_k(x; \lambda) = \lambda_1 f_1(x) + \lambda_2 f_2(x) + \ldots + \lambda_k f_k(x)$$

are defined by the conditions

$$f(x_r) - F_k(x_r; \lambda) = 0 \qquad (r = 1, 2, \ldots, k)$$

where $x_r (r = 1, 2, 3, \ldots, k)$ are the zeros of the polynomial $Q_{k+1}(x)$. If the difference

$$f(x) - F_k(x; \lambda)$$

changes sign alternately at the points $x_r (r = 1, 2, 3, \ldots, k)$ and at no others, then, among all the polynomials $F_k(x; \mu)$ in the metric space $L\ (a, b)$, the polynomial $F_k(x; \lambda)$ deviates the least from $f(x)$. In that connection,

$$\min_{\mu} \int_a^b |f(x) - F_k(x; \mu)|\, dx = \int_a^b |f(x) - F_k(x; \lambda)|\, dx =$$

$$= \left| \int_a^b f(x) \operatorname{sign} Q_{k+1}(x)\, dx \right|.$$

Proof: By virtue of the relation (1), it follows from

$$\int_a^b |f(x) - F_k(x; \mu)|\, dx \geqq \left| \int_a^b \{f(x) - F_k(x; \mu)\} \operatorname{sign} Q_{k+1}(x)\, dx \right|$$

that

$$\int_a^b |f(x) - F_k(x; \mu)|\, dx \geqq \left| \int_a^b f(x) \operatorname{sign} Q_{k+1}(x)\, dx \right|.$$

On the other hand,

$$\int_a^b |f(x) - F_k(x; \lambda)|\, dx =$$

$$= \left| \int_a^b \{f(x) - F_k(x; \lambda)\} \operatorname{sign} Q_{k+1}(x)\, dx \right| = \left| \int_a^b f(x) \operatorname{sign} Q_{k+1}(x)\, dx \right|.$$

Thus the theorem is proved.

For the application of this theorem, it is necessary that the polynomial $Q_k(x)$ be known. Hence the following statement of fact resulting from our considerations is to be noted: If the polynomial

$$F(x) = f_k(x) + a_1 f_1(x) + \ldots + a_{k-1} f_{k-1}(x)$$

satisfies the equations

$$\int_a^b [\operatorname{sign} F(x)] f_r(x)\, dx = 0 \qquad (r = 1, 2, \ldots, k-1),$$

then $F(x) = Q_k(x)$, that is to say, among all the polynomials of the form

$$f_k(x) + \beta_1 f_1(x) + \ldots + \beta_{k-1} f_{k-1}(x)$$

$F(x)$ deviates the least from zero, in the metric space $L(a, b)$.

51. Special Cases of the Theorem of A. A. Markoff. We will begin with a remark. Suppose the integrable function $F(x)$ satisfies the relation

$$F(x + \pi) = -F(x)$$

(and consequently has the period 2π). Suppose that m, n are integers and that the quotient $\frac{m}{n}$ is not an odd number. Then we shall have

$$\int_{-\pi}^{\pi} e^{imx} F(nx)\, dx = 0.$$

By virtue of the aforementioned periodicity,

$$I \equiv \int_{-\pi}^{\pi} e^{imx} F(nx)\, dx = \int_{-\pi+\alpha}^{\pi+\alpha} e^{imx} F(nx)\, dx;$$

when $\alpha = \frac{\pi}{n}$, we get

$$I = e^{\frac{im\pi}{n}} \int_{-\pi}^{\pi} e^{imx} F(nx + \pi)\, dx = -Ie^{\frac{im\pi}{n}}.$$

For example, if we set $F(x) = \operatorname{sign} \cos x$, we obtain

$$\int_0^{\pi} \cos mx \cdot \operatorname{sign} \cos nx\, dx = 0 \qquad (m = 0, 1, 2, \ldots, n-1). \tag{1}$$

Analogously, we have

$$\int_0^{\pi} \sin mx \cdot \operatorname{sign} \sin nx \, dx = 0 \qquad (m = 1, 2, \ldots, n-1). \tag{2}$$

Let us consider the sequence of functions

$$1, \cos x, \cos 2x, \ldots. \tag{3}$$

They form a Markoff system with respect to the interval $[0, \pi]$. Taking equation (1) into consideration, we conclude, by reason of a remark made at the end of the previous section, that

$$Q_k(x) = \operatorname{sign} \cos (k-1)x \qquad (k = 1, 2, \ldots)$$

Whence it follows: If $f(x)$ is a function continuous in the interval $[0, \pi]$ and if

$$S_{n-1}(x) = a_0 + a_1 \cos x + \ldots + a_{n-1} \cos (n-1)x$$

is a trigonometric sum such that the difference

$$f(x) - S_{n-1}(x)$$

changes sign alternately at the points

$$\frac{(2k-1)\,\pi}{2n} \qquad (k = 1, 2, \ldots, n)$$

and at no others in the interval $(0, \pi)$, then we have:

$$\min_{A_k} \int_0^{\pi} |f(x) - A_0 - A_1 \cos x - \ldots - A_{n-1} \cos (n-1)x| \, dx =$$

$$= \int_0^{\pi} |f(x) - S_{n-1}(x)| \, dx = \left| \int_0^{\pi} f(x) \operatorname{sign} \cos nx \, dx \right| .$$

It is easy to see how the foregoing result must be modified if the functions (3) are replaced by the functions

$$\sin x, \sin 2x, \sin 3x, \ldots$$

which form a Markoff system of functions with respect to the open interval $(0, \pi)$. In the latter case, the minimum problem is easy to solve, provided there exists a trigonometric sum

$$S_n(x) = b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx$$

for which the difference

$$f(x) - S_n(x)$$

changes sign alternately at just those points, inside the interval $(0, \pi)$, where the function $\sin(n+1)x$ does so, namely, at the points $\frac{k\pi}{n+1}$ $(k = 1, 2, \ldots, n)$. Thus we obtain

$$\min_{B_k} \int_0^\pi |f(x) - B_1 \sin x - B_2 \sin 2x - \ldots - B_n \sin nx| \, dx =$$

$$= \int_0^\pi |f(x) - S_n(x)| \, dx = \left| \int_0^\pi f(x) \operatorname{sign} \sin (n+1)x \, dx \right|.$$

Essentially the same case presents itself for the functions

$$1, \; x, \; x^2, \ldots.$$

These functions form a Markoff system with respect to any interval. Let us choose the interval $[-1, 1]$.

We will assume that, for the function $f(x)$, there exists a polynomial

$$P_{n-1}(x) = c_0 + c_1 x + \ldots + c_{n-1} x^{n-1}$$

such that the difference

$$f(x) - P_{n-1}(x)$$

changes sign alternately at the points

$$-\cos \frac{k\pi}{n+1} \qquad (k = 1, 2, \ldots, n)$$

and at no other points of the interval $(-1, 1)$. That is true if, for example, $f^{(n)}(x)$ exists and is never zero in the interval $(-1, 1)$.

Let us set $x = \cos\varphi$ and observe that $\frac{\sin(k+1)\varphi}{\sin\varphi}$ is a polynomial of the k-th degree in x (with 2^k as its leading coefficient); let us represent an arbitrary polynomial in the form

$$R_{n-1}(x) = B_1 \frac{\sin\varphi}{\sin\varphi} + B_2 \frac{\sin 2\varphi}{\sin\varphi} + \ldots + B_n \frac{\sin n\varphi}{\sin\varphi} = \frac{1}{\sin\varphi} \sigma_n(\varphi)$$

We may write

$$\int_{-1}^{1} |f(x) - R_{n-1}(x)| \, dx = \int_0^\pi |\sin\varphi \, f(\cos\varphi) - \sigma_n(\varphi)| \, d\varphi$$

and thus we obtain

$$\min_{B_k} \int_{-1}^{1} |f(x) - R_{n-1}(x)|\, dx = \int_{-1}^{1} |f(x) - P_{n-1}(x)|\, dx =$$

$$= \left| \int_{0}^{\pi} \sin\varphi\, f(\cos\varphi)\, \operatorname{sign} \sin(n+1)\varphi\, d\varphi \right| =$$

$$= \left| \int_{-1}^{1} f(x)\, \operatorname{sign} U_n(x)\, dx \right|$$

where

$$U_n(x) = \frac{\sin(n+1)\varphi}{\sin\varphi} \qquad (x = \cos\varphi)\,.$$

In particular, if

$$f(x) = x^n = \frac{1}{2^n}\frac{\sin(n+1)\varphi}{\sin\varphi} + A\frac{\sin n\varphi}{\sin\varphi} + \ldots,$$

then we have

$$\int_{0}^{\pi} \sin\varphi\, f(\cos\varphi)\, \operatorname{sign} \sin(n+1)\,\varphi\, d\varphi =$$

$$= \int_{0}^{\pi} \sin\varphi \left\{ \frac{1}{2^n}\frac{\sin(n+1)\varphi}{\sin\varphi} + \ldots \right\} \operatorname{sign} \sin(n+1)\,\varphi\, d\varphi =$$

$$= \frac{1}{2^n}\int_{0}^{\pi} \sin(n+1)\,\varphi\, \operatorname{sign} \sin(n+1)\,\varphi\, d\varphi =$$

$$= \frac{1}{2^n}\int_{0}^{\pi} |\sin(n+1)\,\varphi|\, d\varphi = \frac{1}{2^n(n+1)} \int_{0}^{(n+1)\pi} |\sin\psi|\, d\psi =$$

$$= \frac{1}{2^n}\int_{0}^{\pi} \sin\psi\, d\psi = \frac{1}{2^{n-1}}\,.$$

Consequently,

$$\min_{C_k} \int_{-1}^{1} |x^n + C_{n-1}x^{n-1} + \ldots + C_0|\, dx = \frac{1}{2^{n-1}}\,.$$

This result is due to A. N. Korkin and E. I. Solotareff [37] [8]

---

[8] It is implicitly contained in a prior work of Tchebysheff [35].

CHAPTER III

ELEMENTS OF HARMONIC ANALYSIS

52. The Simplest Properties of Fourier Series [38]. Let $f(x)$ $(0 \leq x < 2\pi)$ be an integrable function, and let

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,dx, \quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx\,dx \qquad (k = 0, 1, 2, \ldots)$$

be its Fourier coefficients. It is often advantageous to write the Fourier series

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx),$$

which represents the function $f(x)$ in a purely formal manner, in the so-called complex form

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,e^{-ikx}\,dx \qquad (k = 0, \pm 1, \pm 2, \ldots).$$

If we extend $f(x)$ beyond the limits of the interval $[0, 2\pi]$ by means of the equation

$$f(x + 2\pi) = f(x) \qquad (-\infty < x < \infty),$$

then the Fourier coefficients c_k satisfy the equation

$$c_k = \frac{1}{2\pi}\int_{\alpha}^{2\pi+\alpha} f(x)\,e^{-ikx}\,dx$$

for all real values of α. In what follows, it will always be assumed that $f(x)$ has been extended in the manner indicated above.

The following statements are immediate consequences of the definition of Fourier series:

1. If A, B are (complex) constants from

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx}, \qquad \varphi(x) \sim \sum_{-\infty}^{\infty} \gamma_k e^{ikx}$$

follows

$$Af(x) + B\varphi(x) \sim \sum_{-\infty}^{\infty} (Ac_k + B\gamma_k)e^{ikx}.$$

2. If $f(x)$ can be represented by the uniformly convergent series

$$\sum_{-\infty}^{\infty} c_k e^{ikx},$$

then this series is identical with the Fourier series of $f(x)$.

3. If

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx}$$

and if t is a real number, then we have

$$g(x) = f(x+t) \sim \sum_{-\infty}^{\infty} c_k e^{ikt} e^{ikx}.$$

We call attention to a number of additional properties:

4. If

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx},$$

then the expression

$$s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

— the so-called Fourier sum of the n-th order of the function $f(x)$ — satisfies the equation

$$s_n(x) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \frac{\sin\left(n + \frac{1}{2}\right)(t-x)}{\sin\frac{1}{2}(t-x)}\, dt.$$

Here the limits of integration 0, 2π can be replaced by α, $2\pi + \alpha$, where α may be chosen arbitrarily (real).

5. If $f(x)$ is an absolutely continuous periodic function, then the relation

$$f'(x) \sim \sum_{-\infty}^{\infty} ikc_k e^{ikx}$$

follows from

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx}.$$

6. If

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx},$$

then

$$F(x) = \int \{f(x) - c_0\}\, dx$$

is an absolutely continuous periodic function, and we have

$$F(x) \sim C_0 + \sum_{-\infty}^{\infty}{}' \frac{c_k}{ik} e^{ikx}.$$

(The prime means that the summation is to extend only through all values of $k \neq 0$.)

7. From

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx}, \qquad \varphi(x) \sim \sum_{-\infty}^{\infty} \gamma_k e^{ikx}$$

and

$$f(x) \in L^2, \qquad \varphi(x) \in L^2$$

it follows that

$$\frac{1}{2\pi}\int_0^{2\pi} f(t+x)\,\overline{\varphi(x)}\, dx = \sum_{-\infty}^{\infty} c_k \gamma_k e^{ikt},$$

where the series on the right-hand side is absolutely and uniformly convergent (in t).

8. If the corresponding Fourier coefficients of two (integrable) functions coincide throughout, then these two functions are identical almost everywhere.

9. For every (integrable) function we have $\lim_{|k|\to\infty} c_k = 0$.

We shall give concise proofs of the foregoing statements: No. 4 follows from the identity

$$\sum_{k=-n}^{n} e^{ikv} = \frac{e^{i(n+1)v} - e^{-inv}}{e^{iv} - 1} = \frac{\sin\left(n + \frac{1}{2}\right)v}{\sin\frac{1}{2}v}.$$

In connection with the definition of the Fourier sums, we observe that, under the hypothesis $f(x) \in L^2$, the inequality

$$\int_0^{2\pi} |f(x) - s_n(x)|^2\, dx < \int_0^{2\pi} \left| f(x) - \sum_{k=-n}^{n} A_k e^{ikx} \right|^2 dx$$

is valid, provided that the numbers A_k do not all coincide with the corresponding numbers c_k (cf. Section 16).

No. 5 is a consequence of the identity

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)\, e^{-ikx}\, dx =$$

$$= -\frac{f(x)\, e^{-ikx}}{2\pi ik}\Bigg|_0^{2\pi} + \frac{1}{2\pi ik}\int_0^{2\pi} f'(x) e^{-ikx}\, dx = \frac{1}{ik}\cdot\frac{1}{2\pi}\int_0^{2\pi} f'(x)\, e^{-ikx}\, dx$$

where $k \neq 0$.

No. 6 results from No. 5.

No. 7 follows from Parseval's equation (Section 25) and No. 3.

In order to prove No. 8, the so-called uniqueness theorem, it suffices to show that, from $f(x) \in L\,(0, 2\pi)$ and

$$\int_0^{2\pi} f(x)\, e^{-ikx}\, dx = 0 \qquad (k = 0, \pm 1, \pm 2, \ldots)$$

there follows: $f(x)$ is equal to zero almost everywhere.

Let us set

$$F(x) = \int f(x)\, dx \sim \sum_{-\infty}^{\infty} C_k e^{ikx}.$$

Then, by virtue of No. 6, we have

$$C_k = 0 \qquad (k = \pm 1, \pm 2, \ldots). \tag{1}$$

Now, the function

$$F(x) - C_0$$

is continuous and hence is contained in L^2. Therefore, (with the aid of Parseval's equation), we get from (1) the equation

$$F(x) - C_0 \equiv 0.$$

But the derivative of $F(x)$ is equal to $f(x)$ almost everywhere. This proves statement No. 8.

Hence there remains only statement No. 9, the so-called Riemann-Lebesgue theorem. Pursuant to the theorem in Section 24, let us construct a trigonometric sum $\sigma(x)$ such that

$$\frac{1}{2\pi}\int_0^{2\pi} |f(x) - \sigma(x)|\, dx < \varepsilon$$

where ε is any positive number. Then it follows that

$$\left|\frac{1}{2\pi}\int_0^{2\pi} \{f(t) - \sigma(t)\}\, e^{-ikt}\, dt\right| < \varepsilon.$$

Now, since

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(t)\, e^{-ikt}\, dt = \frac{1}{2\pi}\int_0^{2\pi} \{f(t) - \sigma(t)\}\, e^{-ikt}\, dt + \frac{1}{2\pi}\int_0^{2\pi} \sigma(t)\, e^{-ikt}\, dt$$

and the second term on the right vanishes for $|k| > N$ (where N denotes the order of the trigonometric sum $\sigma(x)$), it follows that

$$|c_k| < \varepsilon$$

for $|k| > N$. This proves statement No. 9.

53. Fourier Series for Functions of Bounded Variation. Let $f(x)$ be a periodic function of a real or a complex variable, and suppose that $f(x)$ is of bounded variation (denoted by V) in a closed interval 2π in length. Suppose, in addition, that at every point x of this interval we have

$$f(x) = \frac{f(x+0) + f(x-0)}{2}.$$

Let

$$\sum_{-\infty}^{\infty} c_k e^{ikx}$$

be the Fourier series for $f(x)$. We will show first that for $k \neq 0$ the estimate

$$|c_k| \leqq \frac{V}{2\pi |k|}$$

is valid. In fact, integration by parts gives

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)\, e^{-ikx}\, dx = \frac{1}{2\pi i k}\int_0^{2\pi} e^{-ikx}\, df(x).$$

Whence it follows that

$$|c_k| \leqq \frac{1}{2\pi |k|}\int_0^{2\pi} |df(x)| = \frac{V}{2\pi |k|}.$$

Let us now consider the Fourier sums of the function $f(x)$, which we will write in the form

$$s_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+t)\, \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{1}{2}t}\, dt.$$

We will prove that

$$|s_n(x)| \leqq M + \lambda V$$

with

$$M = \sup |f(x)|.$$

Here λ is a (positive) constant which is always less than unity. (It can be shown [1] that this constant never exceeds $\frac{1}{\pi}$.)

Integration by parts gives

$$s_n(x) = \frac{1}{2\pi} f(x+t) \int_0^t \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv \Bigg|_{t=-\pi}^{t=\pi} -$$

$$- \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \int_0^t \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv \right\} df(t+x).$$

Whence it follows that

$$|s_n(x)| \leqq M + \frac{1}{2\pi} \max_{0 \leq t \leq \pi} \left| \int_0^t \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv \right| \cdot \int_{-\pi}^{\pi} |df(t)|.$$

If we keep in mind that

$$\frac{1}{2\pi} \max_{0 \leq t \leq \pi} \left| \int_0^t \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv \right| =$$

[1] For the proof, we write $s(x)$ in the form

$$s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx} = \sum_{k=-n}^{n} \left(1 - \frac{|k|}{n}\right) c_k e^{ikx} + \frac{1}{n} \sum_{k=-n}^{n} |k|\, c_k e^{ikx} =$$

$$= \sigma_n(x) + \frac{1}{n} \sum_{k=-n}^{n} |k|\, c_k e^{ikx}.$$

Whence it follows that

$$|s_n(x)| \leqq |\sigma_n(x)| + \frac{1}{n} \cdot \frac{V}{2\pi} \cdot 2n = |\sigma_n(x)| + \frac{1}{n} V.$$

It remains merely to prove that

$$|\sigma_n^*(x)| \leqq M.$$

This inequality will in fact be proved later on (cf. Section 61).

$$= \frac{1}{2\pi} \int_0^{\frac{2\pi}{2n+1}} \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv < \frac{1}{2}\int_0^{\pi} \frac{\sin t}{t}\, dt = 0{,}9259\ldots .$$

our proposition is established.

We now proceed to prove the following important theorem of C. Jordan:

THEOREM: *If $f(x)$ is an integrable function whose variation is bounded in the interval (a, b), then the Fourier series for $f(x)$ converges at every point x_0 of this interval to the value of the function at that point, that is to say $f(x)$ approaches*

$$f(x_0) = \frac{f(x_0+0)+f(x_0-0)}{2}.$$

Proof: For the sake of brevity, we will set

$$\frac{f(x_0+t)+f(x_0-t)}{2} = \psi(t)$$

and consider the difference

$$s_n(x_0) - f(x_0) = \frac{1}{\pi}\int_0^{\pi} \{\psi(t) - \psi(+0)\} \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{1}{2}t}\, dt =$$

$$= \frac{1}{\pi}\int_0^{\frac{\pi}{n+\frac{1}{2}}} + \frac{1}{\pi}\int_{\frac{\pi}{n+\frac{1}{2}}}^{\delta} + \frac{1}{\pi}\int_{\delta}^{\pi} = I_1 + I_2 + I_3,$$

wherein we take

$$\frac{\pi}{n+\frac{1}{2}} < \delta < \pi .$$

By hypothesis, there exists an interval $[0, h]$ $(h > 0)$ in which $\psi(t)$ is of bounded variation. For the sake of simplicity, we may assume that $\psi(t)$ does not decrease in this interval for, as we know, every function of bounded variation can be represented as a linear combination of at most four such functions. If we now preassign any $\varepsilon > 0$, we can find a number $\delta = \delta(\varepsilon)$ such that the estimate

$$\psi(t) - \psi(+0) < \frac{\varepsilon}{4\pi}$$

is valid for $0 < t \leq \delta$. For a δ so chosen, we shall have

$$|I_1| < \frac{1}{\pi} \cdot \frac{\varepsilon}{4\pi} \int_0^{\frac{\pi}{n+\frac{1}{2}}} \frac{\left(n+\frac{1}{2}\right)t}{\frac{1}{\pi}t}\, dt = \frac{\varepsilon}{4}.$$

It follows furthermore that

$$I_2 = \frac{1}{\pi}\{\psi(t) - \psi(+0)\} \int_0^t \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv \Bigg|_{t=\frac{\pi}{n+\frac{1}{2}}}^{t=\delta} -$$

$$- \frac{1}{\pi} \int_{\frac{\pi}{n+\frac{1}{2}}}^{\delta} \left\{ \int_0^t \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv \right\} d\psi(t).$$

Consequently,

$$|I_2| < \frac{3\varepsilon}{4\pi} \cdot \frac{1}{\pi} \cdot \max_{0 \leq t \leq \pi} \left| \int_0^t \frac{\sin\left(n+\frac{1}{2}\right)v}{\sin\frac{1}{2}v}\, dv \right| < \frac{3\varepsilon}{4}.$$

If δ is held fixed, we finally obtain

$$\lim_{n\to\infty} I_3 = 0,$$

inasmuch as I_3 is essentially equal to the constant term of the Fourier series of the integrable function

$$\varphi(t) = \begin{cases} \dfrac{\psi(2t) - \psi(+0)}{\sin t} & \left(\dfrac{\delta}{2} < t < \dfrac{\pi}{2}\right), \\ 0 & \text{elsewhere.} \end{cases}$$

Thus it follows that

$$\overline{\lim_{n\to\infty}}\, |s_n(x_0) - f(x_0)| < \varepsilon.$$

But since $\varepsilon > 0$ is an arbitrary positive number, our theorem is proven.

54. The Parseval Equation for Fourier Series. The equation

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)\,\overline{\varphi(x)}\,dx = \lim_{n\to\infty}\sum_{k=-n}^{n} c_k\bar{\gamma}_k \tag{1}$$

where

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx}, \qquad \varphi(x) \sim \sum_{-\infty}^{\infty} \gamma_k e^{ikx}$$

was previously established (Section 52) for the case in which

$$f(x),\ \varphi(x) \in L^2 .$$

However, the equation is also valid under weaker hypotheses. Young [39] has found the following

THEOREM. *The equation* (1) *is valid on the hypothesis that* $f(x)$ *is within* $L(0, 2\pi)$ *and* $\varphi(x)$ *is of bounded variation.*

Proof: Let

$$\varphi_n(x) = \sum_{k=-n}^{n} \gamma_k e^{ikx}.$$

In Section 53 we proved the existence of a constant A having the property that the estimate

$$|\,\varphi_n(x)\,| \leqq A$$

holds for all values of x and n. Hence, for all values of x and n, we have

$$|f(x)\,\overline{\varphi_n(x)}\,| \leqq A\,|f(x)\,|.$$

Since, however,

$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x)$$

almost everywhere it follows with the aid of Lebesgue's theorem that

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)\,\overline{\varphi(x)}\,dx = \lim_{n\to\infty}\frac{1}{2\pi}\int_0^{2\pi} f(x)\,\overline{\varphi_n(x)}\,dx.$$

Now it suffices to keep in mind that

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)\,\overline{\varphi_n(x)}\,dx = \sum_{k=-n}^{n} c_k\bar{\gamma}_k$$

and our proof is complete.

It will be observed that, as a consequence of the foregoing demonstration, the equation [2]

$$\frac{1}{2\pi}\int_0^{2\pi} f(x+t)\,\overline{\varphi}(x)\,dx = \lim_{n\to\infty}\sum_{k=-n}^{n} c_k\overline{\gamma}_k e^{ikt}$$

holds generally.

55. Examples of Fourier Series. Let us consider a periodic function $\varphi_1(t)$ which is defined in the interval $(-\pi, \pi)$ by the formula

$$\varphi_1(t) = \frac{1}{2}t \ .$$

This function is of bounded variation. Its Fourier coefficients have the form

$$c_0 = 0, \qquad c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{t}{2}e^{-ikt}\,dt = \frac{(-1)^{k-1}}{2ik} \qquad (k \neq 0).$$

Therefore the development

$$\varphi_1(t) = \sum_{k=-\infty}^{\infty}{}' \frac{(-1)^{k-1}}{2ik}e^{ikt}$$

or

$$\varphi_1(t) = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{\sin kt}{k}$$

is valid.

By the process of repeated integration, we now introduce the functions $\varphi_r(t)$ $(r = 2, 3, \ldots)$:

$$\varphi_r(t) = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{\cos\left(kt - \frac{\pi r}{2}\right)}{k^r}.$$

Since

$$\varphi'_{r+1} = \varphi_r$$

we find that $\varphi_r(t)$ is a polynomial of the r-th degree in the interval $(-\pi, \pi)$.

As a second example, let us consider the function

$$\psi_1(t) = \ln\left(2\cos\frac{t}{2}\right) \qquad (-\pi < t < \pi).$$

[2] Sometimes the following formula also proves to be useful:

$$\frac{1}{2\pi}\int_0^{2\pi} f(t-x)\,\varphi(x)\,dx = \lim_{n\to\infty}\sum_{k=-n}^{n} c_k\overline{\gamma}_k e^{ikt}.$$

It is integrable and is of bounded variation in every interval

$$(-\pi+\varepsilon,\ \pi-\varepsilon), \qquad (0<\varepsilon<\pi).$$

For this function we have

$$c_k=\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\left(2\cos\frac{t}{2}\right)e^{-ikt}\,dt=\frac{1}{\pi}\int_{0}^{\pi}\ln\left(2\cos\frac{t}{2}\right)\cos kt\,dt=c_{-k}.$$

By virtue of a familiar formula of the integral calculus, we conclude from the previous equation that $c_0=0$. On the other hand, for $k\neq 0$, integration by parts gives the equation

$$c_k=\frac{1}{2\pi}\int_{0}^{\pi}\frac{\sin kt}{k}\,\frac{\sin\frac{t}{2}}{\cos\frac{t}{2}}\,dt.$$

This formula yields in the first place

$$c_1=\frac{1}{2}$$

and secondly

$$kc_k+(k-1)\,c_{k-1}=\frac{1}{\pi}\int_{0}^{\pi}\sin\left(k-\frac{1}{2}\right)t\sin\frac{t}{2}\,dt=0 \qquad (k=2,3,\ldots).$$

Hence we have

$$kc_k=-(k-1)c_{k-1}=\ldots=(-1)^{k-1}c_1=\frac{(-1)^{k-1}}{2},$$

that is

$$\psi_1(t)=\sum_{k=1}^{\infty}(-1)^{k-1}\frac{\cos kt}{k} \qquad (-\pi<t<\pi).$$

Using the method by which we found the functions $\varphi_r(t)$, we introduce the analogous functions

$$\psi_r(t)=\sum_{k=1}^{\infty}(-1)^k\frac{\sin\left(kt-\frac{\pi r}{2}\right)}{k^r} \qquad (r=2,3,\ldots).$$

We then have

$$\psi'_{r+1}=\psi_r.$$

As additional examples, let the reader verify the following developments:

$$\frac{\pi\operatorname{ch}\mu t}{2\mu\operatorname{sh}\mu\pi}=\frac{1}{2\mu^2}+\sum_{k=1}^{\infty}(-1)^k\frac{\cos kt}{\mu^2+k^2} \qquad (-\pi\leqq t\leqq\pi);$$

$$\frac{\pi \operatorname{sh} \mu t}{2 \operatorname{sh} \mu\pi} = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{k \sin kt}{\mu^2 + k^2} \qquad (-\pi < t < \pi);$$

$$\frac{\pi \cos \lambda t}{2\lambda \sin \lambda\pi} = \frac{1}{2\lambda^2} + \sum_{k=1}^{\infty} (-1)^k \frac{\cos kt}{\lambda^2 - k^2} \qquad (-\pi \leqq t \leqq \pi);$$

$$\frac{\pi \sin \lambda t}{2 \sin \lambda\pi} = \sum_{k=1}^{\infty} (-1)^k \frac{k \sin kt}{\lambda^2 - k^2} \qquad (-\pi < t < \pi),$$

where λ is not an integer.

The first of these formulas shows that

$$\frac{\pi \operatorname{ch} \mu t}{2\mu \operatorname{sh} \mu\pi} = \frac{1}{2} + \mu^2 \varphi_2(t) + \mu^4 \varphi_4(t) + \mu^6 \varphi_6(t) + \ldots$$

Whence we find

$$\frac{\mu \operatorname{sh} \mu t}{2 \operatorname{sh} \mu\pi} = \varphi_1(t) + \mu^2 \varphi_3(t) + \mu^4 \varphi_5(t) + \ldots$$

where $|\mu| < 1$, $-\pi < t < \pi$.

If we combine these developments, we obtain

$$\frac{\pi\mu}{2 \operatorname{sh} \mu\pi} e^{\mu t} = \frac{1}{2} + \mu \varphi_1(t) + \mu^2 \varphi_2(t) + \mu^3 \varphi_3(t) + \ldots.$$

Whence we conclude that the polynomials $\varphi_k(t)$ are essentially the same as Bernoulli's polynomials. The latter are defined by the development [40]

$$\frac{\omega e^{h\omega}}{e^{\omega} - 1} = B_0(h) + \frac{\omega}{1!} B_1(h) + \frac{\omega^2}{2!} B_2(h) + \ldots$$

In fact, for $-\pi \leq t \leq \pi$ we have

$$\varphi_k(t) = \frac{2^{k-1}\pi^k}{k!} B_k\left(\frac{1}{2} + \frac{t}{2\pi}\right) \qquad (k = 1, 2, 3, \ldots).$$

Finally, let us consider the function

$$\operatorname{sign} \sin t,$$

which we have already met in Section 51.

From the equation

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} (\operatorname{sign} \sin t)\, e^{-ikt}\, dt = -\frac{i}{\pi} \int_{0}^{\pi} \sin kt\, dt = \frac{i}{k\pi} \{(-1)^k - 1\}$$

there follows

$$\operatorname{sign} \sin t = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\sin (2k+1)t}{2k+1}.$$

Substituting $\frac{\pi}{2} - t$ for t, we get

$$\operatorname{sign} \cos t = \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{\cos (2k+1)t}{2k+1}.$$

56. Trigonometric Integrals [41]. Integrals of the form

$$(1) \qquad \int_{-\infty}^{\infty} C(u)\, e^{iux}\, du, \qquad \int_{0}^{\infty} A(u) \cos ux\, du, \qquad \int_{0}^{\infty} B(u) \sin ux\, du,$$

where x is a real variable, are called trigonometric integrals. Naturally we must indicate the sense in which the integrals are to be understood.

The simplest case occurs if $C(u) \in L(-\infty, \infty)$ or $A(u)$, $B(u) \in L(0, \infty)$. In this case the integrals (1) are proper Lebesgue integrals and, as can be readily seen, represent continuous functions of x.

We shall have under consideration a general case if we merely assume that, for any finite a, b $(a < b)$ and $C(u) \in L(a, b)$, the integral

$$f_{a,b}(x) = \int_{a}^{b} C(u)\, e^{iux}\, du$$

approaches a limit as $a \to -\infty$, $b \to +\infty$. This limit is then an improper integral in the sense of Cauchy and is denoted likewise by the symbol

$$(2) \qquad \int_{-\infty}^{\infty} C(u)\, e^{iux}\, du\,.$$

However, in those cases in which it is impossible to use the fact that the integral is to be understood as a proper Lebesgue integral, it is customary to denote the above-mentioned limit by the symbol

$$(3) \qquad \int_{\to -\infty}^{\to \infty} C(u)\, e^{iux}\, du\,.$$

But it might happen that $f_{a,b}(x)$ has in fact no ordinary limit, and yet a function $f(x)$ of $L^2(-\infty, \infty)$ exists such that

$$\lim_{\substack{a \to -\infty \\ b \to \infty}} \int_{-\infty}^{\infty} |f(x) - f_{a,b}(x)|^2\, dx = 0\,.$$

In this case, the function $f(x)$ is called the limit in the mean [3] (limes in medio) of the functions $f_{a,b}(x)$. Naturally $f(x)$ is determined except for values on a set of measure zero. $f(x)$ is also represented by the integral sign (2). However, in the event that ambiguity is possible, we employ the notation

$$f(x) = \text{l. i. m.}_{\substack{a \to -\infty \\ b \to \infty}} \int_a^b C(u)\, e^{iux}\, du\,. \tag{4}$$

The integrals

$$\int_0^{\to\infty} A(u) \cos ux\, du\,, \qquad \int_0^{\to\infty} B(u) \sin ux\, du\,,$$

$$\text{l. i. m.}_{b\to\infty} \int_0^b A(u) \cos ux\, du\,, \qquad \text{l. i. m.}_{b\to\infty} \int_0^b B(u) \sin ux\, du$$

are defined similarly. They can readily be reduced to the integrals (3), (4).

Later on, we shall see that it is useful to attach the factor $\frac{1}{\sqrt{2\pi}}$ to the trigonometric integral taken over the entire real axis (and the factor $\sqrt{\frac{2}{\pi}}$ to the integral taken over the semi-axes), in other words, to consider, instead of (2), the integral

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u)\, e^{-iux}\, du\,.$$

The function defined by this integral for all, or nearly all, values of x is called the Fourier transform of $g(x)$. We write

$$h(x) = \mathfrak{F}g(x) \quad \text{oder} \quad h = \mathfrak{F}g.$$

If $g_1(x)$ and $g_2(x)$ belong to $L(-\infty, \infty)$ and a_1, a_2 are two arbitrary constants, we have

$$\mathfrak{F}(a_1 g_1 + a_2 g_2) = a_1 \mathfrak{F}g_1 + a_2 \mathfrak{F}g_2. \tag{5}$$

We could mention also other linear manifolds of functions whose elements have Fourier transforms and for which equation (5) is satisfied.

This equation expresses the linear character (homogeneity and additiveness) of the Fourier operator on such manifolds. In that connection, it must be remem-

[3] The limit in the mean can be formed also if $L^2(-\infty, \infty)$ is replaced by the space $L^p(-\infty, \infty)$, where p has any value ≥ 1.

bered that the operator $\mathfrak{F}$ is defined by different constructions for the different types of manifolds, each construction depending on the type of manifold involved (e.g. for $L(-\infty, \infty)$ by the proper Lebesgue integral). Accordingly, the sets corresponding thereto, on which the operator $\mathfrak{F}$ transforms the elements of one manifold or another, will likewise be entirely distinct.

57. The Riemann-Lebesgue Theorem asserts [4]: *The Fourier transform of a function within* $L(-\infty, \infty)$ *converges to zero as* $x \to \pm\infty$.

Let $g(x) \in L(-\infty, \infty)$ and let

$$h(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u)\, e^{-iux}\, du.$$

Let $\varepsilon > 0$ be chosen arbitrarily. Then, corresponding to ε, there exists a positive number A such that

$$\frac{1}{\sqrt{2\pi}} \int_{|u|>A} |\, g(u)\,|\, du < \frac{\varepsilon}{2}\ .$$

Now we have

$$\left| h(x) - \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} g(u)\, e^{-iux}\, du \right| < \frac{\varepsilon}{2} \qquad (-\infty < x < \infty).$$

Furthermore, with the aid of a theorem in Section 24, we can construct a polynomial $P(x)$ so that

$$\frac{1}{\sqrt{2\pi}} \int_{-A}^{A} |\, g(u) - P(u)\,|\, du < \frac{\varepsilon}{2}\ .$$

Consequently,

$$\left| h(x) - \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} P(u)\, e^{-iux}\, du \right| < \varepsilon \qquad (-\infty < x < \infty).$$

It remains merely to show that the function

$$\frac{1}{\sqrt{2\pi}} \int_{-A}^{A} P(u)\, e^{-iux}\, du$$

converges to zero when $x \to \pm\infty$. That can be accomplished at once by integrating by parts:

$$\int_{-A}^{A} P(u)\, e^{-iux}\, du = \frac{P(-A)\, e^{iAx} - P(A)\, e^{-iAx}}{ix} + \frac{1}{ix} \int_{-A}^{A} P'(u)\, e^{-iux}\, du.$$

[4] We have already met a special case of this theorem in Section 52.

58. Plancherel's Theory. Let us now consider the space $L^2(-\infty, \infty)$. If $g(x)$ is situated in $L^2(-\infty, \infty)$, then $g(x)$ belongs also to $L(a, b)$ for all finite values of a, b $(a < b)$. Therefore the integral

$$\frac{1}{\sqrt{2\pi}} \int_a^b g(u)\, e^{-iux}\, du = h_{a,b}(x)$$

exists for all finite values of a, b.

Plancherel [42] was the first to prove that this integral possesses a limit in the mean and thus he was enabled to construct the Fourier operator in the space $L^2(-\infty, \infty)$.

Let us set

$$h(x) = \frac{1}{\sqrt{2\pi}} \operatorname*{l.\,i.\,m.}_{\substack{a \to -\infty \\ b \to \infty}} \int_a^b g(u)\, e^{-iux}\, du. \tag{1}$$

In Plancherel's theory, it is asserted that

$$g(x) = \frac{1}{\sqrt{2\pi}} \operatorname*{l.\,i.\,m.}_{\substack{a \to -\infty \\ b \to \infty}} \int_a^b h(u)\, e^{iux}\, du \tag{2}$$

and, further, that

$$\int_{-\infty}^{\infty} |g(x)|^2\, dx = \int_{-\infty}^{\infty} |h(x)|^2\, dx\ . \tag{3}$$

The above formulas show that the introduction of the factor $\frac{1}{\sqrt{2\pi}}$ is really useful. The foregoing formulas (1), (2) are mutually reciprocal. It is therefore natural to put (1) in the form

$$h = \mathfrak{F} g$$

and to write (2) thus:

$$g = \mathfrak{F}^{-1} h.$$

Hence the operator $\mathfrak{F}$ differs from its inverse $\mathfrak{F}^{-1}$ only as to the sign of the exponent in e^{iux}.

If we conceive of formula (2) as a continuous analogue of the Fourier series, then we may consider (1) as the analogue of the representation of the Fourier coefficients by means of the function to be developed. The relation (3) is therefore obviously an analogue of Parseval's equation and indeed is usually called Parseval's equation.

As we see, there obtains in Plancherel's theory a complete symmetry between a function developed as a trigonometric integral and the Fourier transform of that function (continuous analogue for the series of Fourier coefficients).

Furthermore, the following should be noted: If the Fourier transform of a function within $L^2(-\infty, \infty)$ vanishes almost everywhere, then the function itself is equal to zero almost everywhere.

From the Parseval equation (3), we can derive the generalized Parseval equation for pairs of functions by application of a method previously employed (cf. Section 18). We state here the result without proof:

Let $g_1 \in L^2(-\infty, \infty)$, $g_2 \in L^2(-\infty, \infty)$ and

$$h_1 = \mathfrak{F} g_1, \quad h_2 = \mathfrak{F} g_2;$$

then we have

$$\int_{-\infty}^{\infty} g_1(x)\, \overline{g_2(x)}\, dx = \int_{-\infty}^{\infty} h_1(x)\, \overline{h_2(x)}\, dx. \tag{4}$$

This equation is capable of being further generalized, as follows: From $g(x) = g_2(t - x)$, where t is any constant, we readily obtain

$$h(x) = \mathfrak{F} g(x) = e^{-itx}\overline{h_2(x)}. \tag{4a}$$

If we now apply the equation (4a) to the pair of functions g_1, g_2, we find

$$\int_{-\infty}^{\infty} h_1(x)\, h_2(x)\, e^{itx}\, dx = \int_{-\infty}^{\infty} g_1(x)\, g_2(t - x)\, dx. \tag{5}$$

The right-hand side of this formula is called the *folded* [5] function derived from g_1 and g_2. It is easy to see that this function is continuous with respect to t and converges to zero as $t \to \pm\infty$.

Let us consider $L^2(-\infty, \infty)$ as the Hilbert space in which the product of the elements φ, ψ is defined by the formula

$$(\varphi, \psi) = \int_{-\infty}^{\infty} \varphi(x)\, \overline{\psi(x)}\, dx \ .$$

Therefore the relation (4a) may be written in the form

$$(g_1, g_2) = (\mathfrak{F} g_1, \mathfrak{F} g_2) \ .$$

A homogeneous, additive operator U, defined in all Hilbert space H, is called unitary if the equation

$$(Uf, Uf) = (f, f)$$

holds for any f within H.

Accordingly, the above statement of Plancherel's theory says nothing else but that the Fourier operator is unitary in the space $L^2(-\infty, \infty)$.

[5] Translator's Note: Sometimes called *convoluted* function.

Watson succeeded in obtaining a simple and interesting generalization of Plancherel's theory. In Section 59 we shall give Watson's theorem and its proof. From Watson's theorem, we will then derive Plancherel's theorem.

59. Watson's Theorem [43] reads: *Let the function $\omega(x)$ be defined for all values of x of the form uv, with $a < u < b$, $a < v < b$, where $a \geqq -\infty$, $b \leqq \infty$. In addition, let*

$$\int_a^b \frac{\omega(\lambda t)\,\omega(\mu t)}{t^2}\,dt = \begin{cases} \min\{|\lambda|, |\mu|\} & \text{for } \lambda\mu \geqq 0, \\ 0 & \text{for } \lambda\mu \leqq 0. \end{cases} \tag{1}$$

$$(a < \lambda,\ \mu < b)\ .$$

Then, if $g(x)$ is any function within $L^2(a, b)$, it follows that the function

$$h(x) = \frac{d}{dx}\int_a^b \frac{\omega(xt)}{t}\,g(t)\,dt \tag{2}$$

exists for almost all values of x in the interval (a, b) and is situated in $L^2(a, b)$. Furthermore, we shall have, almost everywhere, the equations:

$$g(x) = \frac{d}{dx}\int_a^b \frac{\omega(xt)}{t}\,h(t)\,dt \tag{3}$$

and

$$\int_a^b |h(x)|^2\,dx = \int_a^b |g(x)|^2\,dx. \tag{4}$$

In other words, the operator defined in $L^2(a, b)$ by formula (2) is unitary.

Titchmarsh [44] has given a simple proof of Watson's theorem; this proof will be repeated here. Without loss of generality we may assume $a < 0$, $b > 0$.

Let us denote by $\mathfrak{M}$ the set of all functions which have a continuous derivative and vanish in every neighborhood about the origin as well as in every neighborhood about each of the points $x = a$, $x = b$.

Suppose, now, that $g(x)$ belongs to $\mathfrak{M}$. Then the function

$$\int_a^b \frac{\omega(xt)}{t}\,g(t)\,dt = \int_{ax}^{bx} \frac{\omega(u)}{u}\,g\left(\frac{u}{x}\right)du \qquad (x \neq 0)$$

has the derivative

$$h(x) = \frac{d}{dx}\int_a^b \frac{\omega(xt)}{t}\,g(t)\,dt = -\frac{1}{x^2}\int_{ax}^{bx} \omega(u)\,g'\left(\frac{u}{x}\right)du = -\frac{1}{x}\int_a^b \omega(xt)\,g'(t)\,dt. \tag{5}$$

By a simple calculation, it follows therefrom that

$$\int_a^b |h(x)|^2\,dx = \int_a^b \frac{dx}{x^2} \int_a^b \omega(xu)\,g'(u)\,du \int_a^b \overline{\omega(xv)}\,\overline{g'(v)}\,dv = \int_a^b |g(t)|^2\,dt.$$

We obtain further — again by virtue of (5) — that

$$\int_a^b \frac{\omega(xt)}{t} h(t)\,dt = -\int_a^b \frac{\omega(xt)}{t^2}\,dt \int_a^b \omega(tu)\,g'(u)\,du =$$

$$= \begin{cases} -\int_a^0 g'(u) \min\{-u, -x\}\,du & \text{for } x \leqq 0, \\ -\int_0^b g'(u) \min\{u, x\}\,du & \text{for } x \geqq 0. \end{cases}$$

Consequently,

$$\int_a^b \frac{\omega(xt)}{t} h(t)\,dt = \int_0^x g(u)\,du,$$

which, however, implies that

$$g(x) = \frac{d}{dx} \int_a^b \frac{\omega(xt)}{t} h(t)\,dt.$$

This proves Watson's theorem for the case where $g(x)$ lies in $\mathfrak{M}$. Let us now observe that the set $\mathfrak{M}$ lies dense in $L^2(a, b)$. That means: For every function $g(x)$ there exists a sequence $\{g_n(x)\} \subset \mathfrak{M}$ such that

$$\lim_{n\to\infty} \int_a^b |g(x) - g_n(x)|^2\,dx = 0\,.$$

It follows from what has already been established that, to every function $g_n(x)$ there corresponds a function $h_n(x)$ which satisfies relations of the form (2), (3), (4). By virtue of (4), we have

$$\int_a^b |h_n(x) - h_m(x)|^2\,dx = \int_a^b |g_n(x) - g_m(x)|^2\,dx.$$

Therefore the sequence $\{h_n(x)\}$ converges in the mean to a certain function $h(x)$ in $L^2(a, b)$, and we thus arrive at the equality

$$\int_a^b |h(x)|^2\,dx = \int_a^b |g(x)|^2\,dx.$$

When x is held fixed, the expressions

$$\int_0^x f(u)\,du, \quad \int_a^b \frac{\omega(xt)}{t} f(t)\,dt, \quad \int_a^b \overline{\frac{\omega(xt)}{t}} f(t)\,dt \qquad (f \in L^2(a,b))$$

are linear functionals in $L^2(a,b)$ and hence are continuous in the sense of continuity of linear functionals. Therefore we have the equations

$$\int_0^x h(u)\,du = \lim_{n\to\infty} \int_0^x h_n(u)\,du = \lim_{n\to\infty} \int_a^b \frac{\omega(xt)}{t} g_n(t)\,dt = \int_a^b \overline{\frac{\omega(xt)}{t}}\, g(t)\,dt$$

and

$$\int_0^x g(u)\,du = \lim_{n\to\infty} \int_0^x g_n(u)\,du = \lim_{n\to\infty} \int_a^b \overline{\frac{\omega(xt)}{t}}\, h_n(t)\,dt = \int_a^b \overline{\frac{\omega(xt)}{t}}\, h(t)\,dt.$$

This completes the proof of Watson's theorem.

60. Plancherel's Theorem. If $a = -\infty$, $b = \infty$, we can substitute for $\omega(x)$ of the previous section the function

$$\frac{1}{\sqrt{2\pi}} \frac{e^{-ix}-1}{-t}.$$

In fact, we have

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{(e^{-i\lambda t}-1)\,(e^{i\mu t}-1)}{t^2}\,dt =$$

$$= \frac{1}{2\pi}\int_{\infty}^{-\infty} \frac{\cos(\mu-\lambda)t - \cos\lambda t - \cos\mu t + 1}{t^2}\,dt =$$

$$= \frac{1}{2\pi}\{|\lambda| + |\mu| - |\mu-\lambda|\}\int_{-\infty}^{\infty} \frac{\sin^2 u}{u^2}\,du = \frac{1}{2}\{|\lambda| + |\mu| - |\mu-\lambda|\} =$$

$$= \begin{cases} \min\{|\lambda|, |\mu|\} & \text{for } \lambda\mu \geqq 0, \\ 0 & \text{for } \lambda\mu \leqq 0. \end{cases}$$

From the general theorem of Watson, there follows Plancherel's theorem: *The operator defined in the space* $L^2(-\infty, \infty)$ *by the formula*

$$h(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \int_{-\infty}^{\infty} \frac{e^{-iux}-1}{-iu}\, g(u)\,du, \tag{1}$$

is unitary. The inverse operator has the form

$$g(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \int_{-\infty}^{\infty} \frac{e^{iux} - 1}{iu} h(u)\, du. \tag{2}$$

We will now show that formulas (1) and (2) can be put in the following form:

$$h(x) = \frac{1}{\sqrt{2\pi}} \operatorname*{l.i.m.}_{n\to\infty} \int_{-n}^{n} g(u)\, e^{-iux}\, du = \mathfrak{F}\, g(x),$$

$$g(x) = \frac{1}{\sqrt{2\pi}} \operatorname*{l.i.m.}_{n\to\infty} \int_{-n}^{n} h(u)\, e^{iux}\, du = \mathfrak{F}^{-1} h(x).$$

It will suffice if we establish the first of these formulas. To that end, let us set

$$g_n(x) = \begin{cases} g(x) & (|x| \leqq n), \\ 0 & (|x| > n). \end{cases}$$

Then $g_n(x)$ lies not only in $L^2(-\infty, \infty)$ but also in $L(-\infty, \infty)$. Therefore we evidently have

$$h_n(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \int_{-\infty}^{\infty} \frac{e^{-iux} - 1}{-iu} g_n(u)\, du = \frac{1}{\sqrt{2\pi}} \int_{-n}^{n} g(u)\, e^{-iux}\, du.$$

Since, however, the equation

$$\int_{-\infty}^{\infty} |h(x) - h_n(x)|^2\, dx = \int_{-\infty}^{\infty} |g(x) - g_n(x)|^2\, dx = \int_{|x|>n} |g(x)|^2\, dx$$

is valid by reason of what we have proved hereinabove, and its extreme right-hand member approaches zero as $n \to \infty$, we have

$$h(x) = \operatorname*{l.i.m.}_{n\to\infty} h_n(x).$$

Thus our proposition is proved.

As exercises, let the reader prove that the two operators $\mathfrak{F}_c$, $\mathfrak{F}_s$ defined in space $L(0, \infty)$ by the formulas

$$\mathfrak{F}_c g(x) = \sqrt{\frac{2}{\pi}} \frac{d}{dx} \int_{0}^{\infty} \frac{\sin ux}{u} g(u)\, du = \sqrt{\frac{2}{\pi}} \operatorname*{l.i.m.}_{n\to\infty} \int_{0}^{n} g(u) \cos ux\, du,$$

$$\mathfrak{F}_s g(x) = \sqrt{\frac{2}{\pi}} \frac{d}{dx} \int_{0}^{\infty} \frac{1 - \cos ux}{u} g(u)\, du = \sqrt{\frac{2}{\pi}} \operatorname*{l.i.m.}_{n\to\infty} \int_{0}^{n} g(u) \sin ux\, du$$

are unitary, and that we have

$$\mathfrak{F}_c^{-1} = \mathfrak{F}_c, \qquad \mathfrak{F}_s^{-1} = \mathfrak{F}_s.$$

61. Fejér's Theorem. According to Weierstrass' second theorem, there exists for every continuous function $f(t)$ of period 2π a sequence of trigonometric sums which converges to $f(t)$ uniformly. Among those trigonometric sums which can possibly be paired with the function $f(t)$, the Fourier sums $s_n(t)$ are doubtless by far the most important. However, even the continuity of $f(t)$ is still not sufficient to guarantee that $s_n(t)$ will converge to $f(t)$ as $n \to \infty$. In that connection, the following discovery of Fejér [45] is of the utmost significance. He proved that the trigonometric sum

$$\sigma_n(t) = \frac{s_0(t) + s_1(t) + \ldots + s_{n-1}(t)}{n},$$

i.e., the arithmetic mean of the first n Fourier sums of the function $f(t)$, converges to $f(t)$ uniformly as $n \to \infty$, if $f(t)$ is a continuous function of period 2π. [6] For that reason, the sums σ_n are called Fejér sums.

Without requiring that $f(t)$ be continuous at this time, let us set

$$f(t) \sim \sum_{-\infty}^{\infty} c_k e^{ikt}.$$

A simple calculation then leads to the relation

$$\sigma_n(t) = \sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) c_k e^{ikt}; \tag{1}$$

since, on the other hand, the equations

$$s_k(t) = \frac{1}{2\pi} \int_0^{2\pi} f(t+u) \frac{\sin\left(k + \frac{1}{2}\right)u}{\sin \frac{1}{2}u}\, du$$

and

$$\sum_{k=0}^{n-1} \frac{\sin\left(k + \frac{1}{2}\right)u}{\sin \frac{1}{2}u} = \sum_{k=0}^{n-1} \frac{\cos ku - \cos(k+1)u}{2 \sin^2 \frac{1}{2}u} = \frac{1 - \cos nu}{2 \sin^2 \frac{1}{2}u} = \frac{\sin^2 \frac{nu}{2}}{\sin^2 \frac{u}{2}}$$

are valid, it follows that

$$\sigma_n(t) = \int_0^{2\pi} f(t+u)\, \Phi_n(u)\, du \tag{2}$$

[6] This gives us also another proof of Weierstrass' theorem.

where

$$\Phi_n(u) = \frac{1}{2\pi n} \frac{\sin^2 \frac{nu}{2}}{\sin^2 \frac{u}{2}}.$$

The function Φ_n is called the *Fejér kernel.* An important property of the Fejér kernel is the fact that it is positive. Furthermore: If $f(t) \equiv 1$, then $s_n(t) \equiv 1$ $(n = 0, 1, 2, \ldots)$, that is, $\sigma_n(t) \equiv 1$ $(n = 1, 2, \ldots)$. Consequently, we have

$$\int_0^{2\pi} \Phi_n(u)\, du = 1. \tag{2'}$$

In addition to showing that Fejér's kernel is positive, this equation shows that the inequality

$$\inf_{0 \leqq t \leqq 2\pi} f(t) \leqq \sigma_n(t) \leqq \sup_{0 \leqq t \leqq 2\pi} f(t)$$

and more generally

$$|\sigma_n(t)| \leqq \sup_{0 \leqq t \leqq 2\pi} |f(t)|$$

holds for real-valued functions $f(t)$.

De la Vallée-Poussin [46] has brought Fejér's sums into a different form. This can be done by means of the development

$$\frac{1}{\sin^2 z} = \sum_{k=-\infty}^{\infty} \frac{1}{(z + \pi k)^2}.$$

In fact, it follows from this development that

$$\Phi_n(u) = \frac{1}{2\pi n} \frac{\sin^2 \frac{nu}{2}}{\sin^2 \frac{u}{2}} = \frac{2}{\pi n} \sum_{k=-\infty}^{\infty} \frac{\sin^2 \frac{nu}{2}}{(u + 2k\pi)^2}$$

where the last series converges uniformly along the entire real axis $-\infty\, u < \infty$. Now, since

$$\sigma_n(t) = \int_{-\pi}^{\pi} f(t + u)\, \Phi_n(u)\, du,$$

there follows

$$\sigma_n(t) = \frac{2}{\pi n} \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} f(t+u) \frac{\sin^2 \frac{nu}{2}}{(u+2k\pi)^2} du =$$

$$= \frac{2}{\pi n} \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} f(t+u+2k\pi) \frac{\sin^2 \frac{n(u+2k\pi)}{2}}{(u+2k\pi)^2} du =$$

$$= \frac{2}{\pi n} \sum_{k=-\infty}^{\infty} \int_{(2k-1)\pi}^{(2k+1)\pi} f(t+x) \frac{\sin^2 \frac{nx}{2}}{x^2} dx = \frac{2}{\pi n} \int_{-\infty}^{\infty} f(t+x) \frac{\sin^2 \frac{nx}{2}}{x^2} dx;$$

whence

$$\sigma_n(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} f\left(t + \frac{2u}{n}\right) \frac{\sin^2 u}{u^2} du. \tag{3}$$

It will be observed that by formula (2′), we have

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2 u}{u^2} du = 1. \tag{3′}$$ [7]

The integral

$$\frac{1}{\pi} \int_{-\infty}^{\infty} f\left(t + \frac{2u}{\lambda}\right) \frac{\sin^2 u}{u^2} du = \frac{1}{\pi} \int_{-\infty}^{\infty} f(v) \frac{2 \sin^2 \frac{\lambda}{2}(t-v)}{\lambda (t-v)^2} dv \tag{4}$$

has a meaning also under more general hypotheses than those given above. In the first place, we do not need the assumption that $f(t)$ has the period 2π. It suffices to assume that the function $f(t)$ $(-\infty < t < \infty)$ is measurable and that

$$\frac{f(t)}{1+t^2} \in L(-\infty, \infty)$$

lies in $L^2(-\infty, \infty)$.

Secondly, it is not necessary to assume that λ is a positive integer. Let us agree that the integral (4), subject to these general assumptions regarding $f(t)$, shall be known as *Fejér's integral* of $f(t)$. The integral (4) is also referred to as the *Fejér operator*. The expression

$$\frac{2}{\pi\lambda} \frac{\sin^2 \frac{\lambda}{2} t}{t^2}$$

is called the *Fejér kernel*.

[7] We have already made use of formula (3′), in Section 60.

The question now arises as to the behavior of the Fejér integral when we pass to the limit $\lambda \to \infty$. Section 62 will be dedicated to the investigation of this question, and we shall be led to a number of general results which contain, in particular, Fejér's theorem for periodic functions, which was mentioned at the beginning of this section.

62. Integral-Operators with Kernels of the Fejér Type. Instead of Fejér's operator, let us consider the more general operator

$$K_\lambda f(x) = \lambda \int_{-\infty}^{\infty} f(u)\, K(\lambda (x-u))\, du \qquad (\lambda > 0) \tag{1}$$

containing the kernel $K(x)$. This kernel is said to be of the *Fejér type* if

a) $K(-x) = K(x)$,

b) $\int_{-\infty}^{\infty} K(x)\, dx = 1$,

c) $K(x)$ is bounded on the interval $[-1, 1]$,

d) $x^2 K(x)$ is bounded along the entire real axis.

It is clear that, when the conditions a)—d) are fulfilled for all positive values of λ, the integral (1) will exist if $f(x)$ is a measurable function $(-\infty < x < \infty)$ which satisfies the condition

$$\frac{f(x)}{1+x^2} \in L(-\infty, \infty)\,. \tag{2}$$

THEOREM. *If $f(x)$ satisfies the condition* (2) *and if*

$$f(x; \lambda) = K_\lambda f(x),$$

where K_λ is an operator with a kernel of the Fejér type, then:

1. *From*

$$\lim_{h \to 0} \frac{1}{h} \int_0^h \left| \frac{f(x+t) + f(x-t)}{2} - s \right| dt = 0 \qquad (h > 0) \tag{3}$$

it follows that

$$\lim_{\lambda \to \infty} f(x; \lambda) = s\,.$$

2. *If $f(x)$ is continuous in the finite interval (α, β), then*

$$\lim_{\lambda \to \infty} f(x; \lambda) = f(x)$$

uniformly in every closed partial interval of (α, β).

Proof. By virtue of the properties a) and b) of the functions $K(x)$, we have for an arbitrary s:

$$f(x;\lambda) - s = \lambda \int_{-\infty}^{\infty} \{f(u) - s\}\, K(\lambda(x-u))\, du =$$

$$= 2\lambda \int_{0}^{\infty} \left\{ \frac{f(x+t)+f(x-t)}{2} - s \right\} K(\lambda t)\, dt.$$

We will now choose $\delta > 0$, $\lambda > \frac{1}{\delta}$ and bring the foregoing expression into the form

$$f(x;\lambda) - s = 2\lambda \int_{0}^{\frac{1}{\lambda}} \left\{ \frac{f(x+t)+f(x-t)}{2} - s \right\} K(\lambda t)\, dt +$$

$$+ 2\lambda \int_{\frac{1}{\lambda}}^{\delta} + 2\lambda \int_{\delta}^{\infty} = I_1 + I_2 + I_3.$$

In conformity with the conditions c), d), let

$$\sup_{-1 \leqq x \leqq 1} |K(x)| = A, \qquad \sup_{-\infty < x < \infty} x^2 |K(x)| = B.$$

Furthermore, let us set

$$\varphi(t) = \varphi(t; x, s) = \frac{1}{t} \int_{0}^{t} \left| \frac{f(x+u)+f(x-u)}{2} - s \right| du.$$

Then we shall have

$$\text{(4)} \qquad \begin{cases} |I_1| \leqq 2A\varphi\left(\frac{1}{\lambda}\right), \\ |I_2| \leqq \frac{2B}{\lambda} \int_{\frac{1}{\lambda}}^{\delta} \left| \frac{f(x+t)+f(x-t)}{2} - s \right| \frac{dt}{t^2} \leqq \\ \leqq \frac{2B}{\lambda} \frac{\varphi(t)}{t} \Big|_{t=\frac{1}{\lambda}}^{\delta} + \frac{4B}{\lambda} \int_{\frac{1}{\lambda}}^{\delta} \frac{\varphi(t)}{t^2}\, dt \leqq 6B \sup_{\frac{1}{\lambda} \leqq t \leqq \delta} \varphi(t). \end{cases}$$

In addition thereto, we obtain

$$\text{(5)} \qquad |I_3| \leqq \frac{2B}{\lambda} \int_{\delta}^{\infty} \left| \frac{f(x+t)+f(x-t)}{2} - s \right| \frac{dt}{t^2}.$$

We will now preassign an arbitrary positive number ε. If (3) is fulfilled for certain values of x and s, we can choose a $\delta = \delta(\varepsilon)$ such that, when $0 < t \leq \delta$,

$$\varphi(t) < \frac{\varepsilon}{2A + 6B}.$$

With such a δ it follows by reason of (4) that

$$|I_1| + |I_2| < \varepsilon.$$

On the other hand, for a fixed δ, we get from (5)

$$\lim_{\lambda \to \infty} I_3 = 0.$$

Therefore

$$\overline{\lim_{\lambda \to \infty}} |f(x; \lambda) - s| \leqq \varepsilon,$$

whence proposition 1 follows, since, as we know, $\varepsilon > 0$ may be chosen arbitrarily.

Let us now assume that $f(x)$ is continuous in a certain finite interval (α, β). Let Δ be any closed partial interval of this interval. Then a $\delta = \delta(\varepsilon)$ can be so chosen that the estimate

$$\left| \frac{f(x+t) + f(x-t)}{2} - f(x) \right| \leqq \frac{\varepsilon}{2A + 6B}$$

holds for $0 \leq t \leq \delta$ and for any value of x in Δ. Whence it follows that for $0 \leq t \leq \delta$ and any value of x in Δ, we have

$$\varphi(t) \leqq \frac{\varepsilon}{2A + 6B},$$

if we put $s = f(t)$. But since, for a fixed δ,

$$|I_1| + |I_2| \leqq \varepsilon$$

uniformly, for every $x \in \Delta$, it follows that the relation

$$\lim_{\lambda \to \infty} I_3 = 0$$

is likewise satisfied uniformly for every $x \in \Delta$. Hence we have, uniformly in Δ, the inequality

$$\overline{\lim_{\lambda \to \infty}} |f(x; \lambda) - f(x)| \leqq \varepsilon.$$

Thus proposition 2 has been established. As to the relation (3), it is satisfied at all points of continuity of $f(x)$ if we set $s = f(x)$. (3) is also valid, however, at all points of discontinuity of the first kind if we merely set

$$s = \frac{f(x+0) + f(x-0)}{2}.$$

To Lebesgue is due an important theorem, by reason of which the relation (3) holds for nearly all values of x provided that $f(x)$ is integrable in every finite partial interval and that we put $s = f(x)$.

THE THEOREM OF LEBESGUE [48]. *If $f(x) \in L(a, b)$, then we have*

$$\lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} | f(x + t) - f(x) | \, dt = 0$$

almost everywhere in (a, b).

Proof. If we keep a fixed, then, by reason of Lebesgue's theorem concerning the derivative of an indefinite integral, the equation

$$\lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} | f(x + t) - a | \, dt = \lim_{h \to 0} \frac{1}{2h} \int_{x-h}^{x+h} | f(u) - a | \, du = | f(x) - a | \tag{6}$$

is valid almost everywhere in (a, b).

Let us denote by e_a the set of measure zero in which the relation (6) is not satisfied. Hence we have $e_a \subset (a, b)$. Now suppose a runs through the set of all complex rational numbers. This set is denumerable; therefore the union e of all sets e_a is in turn a set of measure zero.

Suppose, now, that x belongs to e. Let us choose any positive number ε, together with a complex rational number β such that

$$| f(x) - \beta | \leqq \frac{\varepsilon}{2} \, .$$

Then it follows that

$$\frac{1}{2h} \int_{-h}^{h} | f(x + t) - f(x) | \, dt \leqq$$

$$\leqq \frac{1}{2h} \int_{-h}^{h} | f(x + t) - \beta | \, dt + \frac{1}{2h} \int_{-h}^{h} | f(x) - \beta | \, dt \leqq \frac{1}{2h} \int_{-h}^{h} | f(x + t) - \beta | \, dt + \frac{\varepsilon}{2}$$

and hence

$$\lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} | f(x + t) - f(x) | \, dt \leqq \varepsilon .$$

Thus the theorem of Lebesgue is proven.

63. The Theorem of Young and Hardy [49]. *Suppose $f(x)$ is an integrable function of period 2π. If $\varphi(x)$ belongs to $L(-\infty, \infty)$ and is of bounded variation along the entire real axis, then the integral*

$$\int_{-\infty}^{+\infty} f(x) \, \varphi(x) \, dx \tag{1}$$

exists and is equal to

$$\lim_{n\to\infty}\sum_{k=-n}^{n} c_k \int_{-\infty}^{\infty} \varphi(x)\, e^{ikx}\, dx, \tag{2}$$

where $c_k (k = 0, \pm 1, \pm 2, \ldots)$ *are the Fourier coefficients of* $f(x)$.

Proof. Let us consider the infinite series

$$\sum_{k=-\infty}^{\infty} \varphi(x + 2k\pi) \tag{3}$$

and let us show that, in the interval $[0, 2\pi]$, this series converges absolutely and uniformly to a function of bounded variation in this interval.

To that end, we set

$$\psi_n(x) = \frac{1}{2\pi}\int_{x+2n\pi}^{x+2(n+1)\pi} \varphi(t)\, dt.$$

Then we have

$$\sum_{n=k}^{k+m} |\psi_n(x)| \leqq \frac{1}{2\pi}\int_{x+2k\pi}^{x+2(k+m+1)\pi} |\varphi(t)|\, dt.$$

But since $\varphi(x) \in L(-\infty, \infty)$, it follows that the series

$$\sum_{-\infty}^{\infty} |\psi_n(x)|$$

converges uniformly in the interval $[0, 2\pi]$. On the other hand,

$$\varphi(x + 2n\pi) - \psi_n(x) = \frac{1}{2\pi}\int_{x+2n\pi}^{x+2(n+1)\pi} \{\varphi(x + 2n\pi) - \varphi(t)\}\, dt,$$

that is,

$$|\varphi(x + 2n\pi)| \leqq |\psi_n(x)| + \operatorname*{Var}_{x+2n\pi \leqq t \leqq x+2(n+1)\pi} \varphi(t) = |\psi_n(x)| + \varepsilon_n(x).$$

By hypothesis, the series

$$\sum_{n=-\infty}^{\infty} \varepsilon_n(x)$$

converges uniformly in every finite partial interval. Consequently, the series (3) converges absolutely and uniformly in the interval $[0, 2\pi]$. We will denote its sum by $\Phi(x)$.

Now, since

$$\sum_j |\Phi(z_j) - \Phi(y_j)| \leqq \sum_{j,k} |\varphi(z_j + 2k\pi) - \varphi(y_j + 2k\pi)|$$

it follows that $\Phi(x)$ is of bounded variation in the interval $[0, 2\pi]$. By applying Young's theorem (cf. Section 54), we find for Parseval's equation the relation:

$$\int_0^{2\pi} f(x)\,\Phi(x)\,dx = \lim_{n\to\infty} \sum_{k=-n}^{n} c_k \int_0^{2\pi} \Phi(x)\,e^{ikx}\,dx.$$

It should be noticed also that, by reason of the uniform convergence of the series (3), the following equations hold:

$$\int_0^{2\pi} f(x)\,\Phi(x)\,dx = \lim_{\substack{m\to-\infty \\ n\to\infty}} \sum_{k=m}^{n} \int_0^{2\pi} f(x)\,\varphi(x+2k\pi)\,dx =$$

$$= \lim_{\substack{m\to-\infty \\ n\to\infty}} \int_{2m\pi}^{2(n+1\,\pi} f(x)\,\varphi(x)\,dx = \int_{-\infty}^{\to\infty} f(x)\,\varphi(x)\,dx,$$

$$\int_0^{2\pi} \Phi(x)\,e^{ikx}\,dx = \int_{-\infty}^{\to\infty} \varphi(x)\,e^{ikx}\,dx = \int_{-\infty}^{\infty} \varphi(x)\,e^{ikx}\,dx.$$

64. Examples of Kernels of the Fejér Type. In various problems of analysis, we encounter special integral-operators with kernels of the Fejér type which are noteworthy in one respect or another. Ordinarily, the function $K(x)$ has, besides the properties enumerated in Section 62, an added property in that its variation is bounded along the whole real axis. Let us now assume this added property to be present, and let us introduce the function

$$\varphi(t) = \int_{-\infty}^{\infty} K(x)\,e^{-itx}\,dx,$$

which is identical with the Fourier transform of $K(x)$ except for a numerical factor.

By reason of the condition

$$\int_{-\infty}^{\infty} K(x)\,dx = 1$$

we have the equation

$$\varphi(0) = 1.$$

Now let us consider the integral

$$f(x;\lambda) = \lambda \int_{-\infty}^{\infty} f(u)\, K(\lambda(x-u))\, du \qquad (\lambda > 0)$$

with

$$\frac{f(x)}{1+x^2} \in L(-\infty, \infty).$$

We will seek to determine the form of the function $f(x;\lambda)$ in each of the two following cases:

a) $f(x)$ has the period 2π;

b) $f(x)$ lies in $L^2(-\infty, \infty)$.

In case a), if we set

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx}$$

and apply the theorem of Young and Hardy (cf. Section 63), we find the relations

$$f(x;\lambda) = \lim_{n\to\infty} \sum_{k=-n}^{n} c_k \lambda \int_{-\infty}^{\infty} K(\lambda(x-u))\, e^{iku}\, du =$$

$$= \lim_{n\to\infty} \sum_{k=-n}^{n} \varphi\left(\frac{k}{\lambda}\right) c_k e^{ikx} = \sum_{k=-\infty}^{\infty} \varphi\left(\frac{k}{\lambda}\right) c_k e^{ikx}$$

$K(x)$	$\varphi(x)$	Kernel
$\frac{1}{\pi}\frac{a}{a^2+x^2}$ $(a>0)$	$e^{-a\,x}$	Abel-Poisson
e^{-ax^2} $(a>0)$	$\frac{1}{2\sqrt{\pi a}}\, e^{-\frac{x^2}{4a}}$	Weierstrass
$\frac{2\sin^2\frac{x}{2}}{\pi x^2}$	$1-\lvert x\rvert$ $(\lvert x\rvert \leqq 1)$ 0 $(\lvert x\rvert \geqq 1)$	Fejér
$\frac{12\sin^4\frac{x}{2}}{\pi x^4}$	$1-\frac{3x^2}{2}+\frac{3\lvert x\rvert^3}{4}$ $(\lvert x\rvert \leqq 1)$ $\frac{1}{4}(2-\lvert x\rvert)^3$ $(1 \leqq \lvert x\rvert \leqq 2)$ 0 $(\lvert x\rvert \geqq 2)$	Jackson — de la Vallée-Poussin

Thus $f(x;\lambda)$ can be obtained from the Fourier series for $f(x)$ by introducing the "factor" $\varphi\left(\frac{k}{\lambda}\right)$.

In case b), we attain a corresponding result. By applying a theorem on the Faltung of functions in L^2 (it is easy to see, from the conditions which $K(x)$ satisfies, that $K(x)$ lies in L^2), we find that

$$f(x;\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi\left(\frac{v}{\lambda}\right) C(v)\, e^{ivx}\, dv,$$

where $C(x)$ is the Fourier transform of the function $f(x)$. Thus the function $f(x;\lambda)$ is obtained by the introduction of a "factor" into the Fourier integral of the function $f(x)$.

A table of noteworthy kernels is given on page 119.

65. The Fourier Transformation of Integrable Functions. As we have already seen, the problem of finding the inverse of the Fourier integral in space $L^2(-\infty, \infty)$ was quite easy to solve. In other spaces, this problem is more complicated. The simplest theorem bearing on this problem [41] states: *If $f(x) \in L(a, b)$ for any finite a, b and the integral*

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\to\infty} f(u)\, e^{iux}\, du \tag{1}$$

converges uniformly in every finite interval of the x-axis, then, almost everywhere and in particular at all points of continuity of $f(x)$, we have

$$f(x) = \frac{1}{\sqrt{2\pi}} \lim_{\lambda\to\infty} \int_{-\lambda}^{\lambda} \left(1 - \frac{|u|}{\lambda}\right) g(u)\, e^{-iux}\, du.$$

Incidentally, it will be observed that the hypothesis of the theorem is surely satisfied if $f(x)$ lies in $L(-\infty, \infty)$.

Proof. We will prove first that, by virtue of our hypothesis, the equation

$$\frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} \left(1 - \frac{|u|}{\lambda}\right) g(u)\, e^{-iux}\, du = \frac{1}{\pi} \int_{-\infty}^{\to\infty} f(t)\, \frac{2\sin^2\frac{\lambda}{2}(x-t)}{\lambda(x-t)^2}\, dt \tag{2}$$

follows from (1). In fact,

$$\frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} \left(1 - \frac{|u|}{\lambda}\right) g(u)\, e^{-iux}\, du = \frac{1}{2\pi} \int_{-\lambda}^{\lambda} \left(1 - \frac{|u|}{\lambda}\right) e^{-iux}\, du \int_{-\infty}^{\to\infty} f(t)\, e^{iut}\, dt =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\to -\infty} f(t)\, dt \int_{-\lambda}^{\lambda} \left(1 - \frac{|u|}{\lambda}\right) e^{iu(t-x)}\, du = \frac{1}{\pi} \int_{-\infty}^{\to\infty} f(t)\, \frac{2\sin^2\frac{\lambda}{2}(x-t)}{\lambda(x-t)^2}\, dt,$$

wherein the interchange in the order of the integration is justified because of the uniform convergence of the integral (1).

Let us now set

$$F(t) = \int_0^t f(u)\,du.$$

Since the integral (1) converges when $x = 0$, it follows that the function $F(t)$ is bounded:

$$|F(t)| \leqq M \qquad (-\infty < t < \infty).$$

Assuming $T > |x|$, we may reduce the right-hand side of the equation (2) to the form

$$\frac{1}{\pi}\int_{-\infty}^{+\infty} f(t)\,\frac{2\sin^2\frac{\lambda}{2}(x-t)}{\lambda(x-t)^2}\,dt = \frac{1}{\pi}\int_{-\infty}^{-T} + \frac{1}{\pi}\int_{-T}^{T} + \frac{1}{\pi}\int_{T}^{+\infty} = I_1 + I_2 + I_3.$$

We are going to show that, when T is sufficiently large (for every $\lambda > \lambda_0$), the integrals I_1, I_3 will be sufficiently small. To that end, we will apply to the integral I_2 a theorem of Section 62; this integral is, of course, the Fejér integral of the function

$$f_T(t) = \begin{cases} f(t) & (|t| \leqq T), \\ 0 & (|t| > T). \end{cases}$$

It suffices to investigate I_3:

$$I_3 = \frac{1}{\pi}\int_{T}^{+\infty} f(t)\,\frac{2\sin^2\frac{\lambda}{2}(x-t)}{\lambda(x-t)^2}\,dt = \frac{1}{\pi}F(t)\,\frac{2\sin^2\frac{\lambda}{2}(x-t)}{\lambda(x-t)^2}\Bigg|_{t=T}^{t=\infty} +$$

$$+\frac{1}{\pi}\int_{T}^{\infty} F(t)\,\frac{\sin\lambda(x-t)}{(x-t)^2}\,dt - \frac{1}{\pi}\int_{T}^{\infty} F(t)\,\frac{4\sin^2\frac{\lambda}{2}(x-t)}{\lambda(x-t)^3}\,dt.$$

Whence it follows that

$$|I_3| \leqq \frac{2M}{\pi\lambda(T-x)^2} + \frac{M}{\pi}\int_{T}^{\infty}\frac{dt}{(t-x)^2} + \frac{4M}{\pi\lambda}\int_{T}^{\infty}\frac{dt}{(t-x)^3} =$$

$$= \frac{4M}{\pi\lambda(T-x)^2} + \frac{M}{\pi(T-x)}.$$

Here the right-hand side can in fact be made as small as we please provided that T is taken sufficiently large, when $\lambda > 1$.

The theorem just proven yields the following elementary "uniqueness theorem": *If $f(x)$ lies in $L(a, b)$, for all finite values of a, b, and if the integral*

$$\int_{\to -\infty}^{\to \infty} f(u)\, e^{iux}\, du = 0$$

converges uniformly in every finite interval on the x-axis, then $f(x)$ is equal to zero almost everywhere.

The theorem proved in this Section admits various generalizations. Let us formulate one of them here: *Suppose $f(x)$ lies in $L(a, b)$, for arbitrary finite values of a, b. Now, if the function*

$$\lim_{\lambda \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} f(u)\, e^{iux}\, du = g(x)$$

is finite everywhere and lies for all finite values of a, b in $L(a, b)$, then, almost everywhere, we have

$$f(x) = \lim_{\lambda \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} \left(1 - \frac{|u|}{\lambda}\right) g(u)\, e^{-iux}\, du.$$

This theorem is due to Offord [50]. A consequence thereof is the following uniqueness theorem [50], [51]: *If, for all finite values of a, b, the function $f(x)$ lies in $L(a, b)$ and if we have*

$$\lim_{\lambda \to \infty} \int_{-\lambda}^{\lambda} f(u)\, e^{iux}\, du = 0$$

for every x, then $f(x) = 0$ almost everywhere.

Offord [50] discovered still another, very remarkable, uniqueness theorem: *If $f(x)$ lies in $L(a, b)$, for arbitrary finite values of a, b, and if*

$$\lim_{\lambda \to \infty} \int_{-\lambda}^{\lambda} \left(1 - \frac{|u|}{\lambda}\right) f(u)\, e^{iux}\, du = 0 \tag{3}$$

for every x, then $f(x)$ is zero almost everywhere.

This theorem is the best possible in the sense that the requirement, the relation (3) shall be valid everywhere, cannot be further weakened. Indeed, the function

$$f(u) = 1$$

satisfies the relation (3) for every $x \neq 0$.

66. The Faltung of Two Functions. In Section 58 (supra), we defined the Faltung of two functions $g_1(x)$, $g_2(x) \in L^2(-\infty, \infty)$ as the integral

$$\int_{-\infty}^{\infty} g_1(x)\, g_2(t - x)\, dx. \tag{1}$$

We saw that this integral is equal to

$$\int_{-\infty}^{\infty} h_1(x)\, h_2(x)\, e^{iux} dx .$$

But, except for the factor $\frac{1}{\sqrt{2\pi}}$, the second integral is the inverse Fourier transform of the function $h_1(x)\, h_2(x) \in L(-\infty, \infty)$. We will agree to denote the function derived from the Faltung of g_1, g_2 by

$$g_1 * g_2 .$$

Then it is clear that

$$g_1 * g_2 = g_2 * g_1 .$$

The result stated at the beginning can now be brought into the form

$$g_1 * g_2 = \sqrt{2\pi}\, \mathfrak{F}^{-1} \{\mathfrak{F} g_1 \cdot \mathfrak{F} g_2$$

We will assume: $g_1(x)$, $g_2(x) \in L(-\infty, \infty)$. Since the integral

$$\int_{-\infty}^{\infty} |\, g_1(x)\, |\, dx \int_{-\infty}^{\infty} |\, g_2(t-x)\, |\, dt$$

exists, it follows from Fubini's theorem that also the integral (1) exists for almost every t and represents a function in $L(-\infty, \infty)$.

Hence, in this case, too, we may speak of a Faltung, which is defined here for almost every x. Its Fourier transform can be found thus:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-itu} dt \int_{-\infty}^{\infty} g_1(x)\, g_2(t-x)\, dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iux}\, g_1(x)\, dx \int_{-\infty}^{\infty} e^{-iu(t-x)}\, g_2(t-x)\, dt = \sqrt{2\pi}\, h_1(u)\, h_2(u) .$$

Consequently, for the space $L(-\infty, \infty)$ we have the relation

$$\mathfrak{F} \{g_1 * g_2\} = \sqrt{2\pi}\, \mathfrak{F} g_1 \cdot \mathfrak{F} g .$$

67. V. A. Stekloff's Functions. Suppose the function $f(t)$, defined along the entire real axis, belongs to $L(a, b)$, for all finite values of a, b. Given any positive h, let us now construct the function

$$f_h(t) = \frac{1}{h} \int_{t-\frac{h}{2}}^{t+\frac{h}{2}} f(u)\, du = \frac{1}{h} \int_{-\frac{h}{2}}^{\frac{h}{2}} f(t+v)\, dv .$$

If $f(t)$ is a periodic function, then $f_h(t)$ will have the same period. Furthermore: According to Lebesgue's theorem, the function $f_h(t)$ has the derivative

$$f_h'(t) = \frac{1}{h}\left\{f\left(t+\frac{h}{2}\right) - f\left(t-\frac{h}{2}\right)\right\}$$

almost everywhere. $f_h(t)$ is called a *Stekloff function* of $f(t)$.

Let us assume that $f(t)$ has the period 2π and that we have

$$f(t) \sim \sum_{-\infty}^{\infty} c_k e^{ikt}.$$

We are going to show how the Fourier series of $f_h(t)$ can be obtained. We already know (Section 52) that

$$F(t) = \int \{f(t) - c_0\}\, dt \sim C_0 + \sum_{-\infty}^{\infty}{}' \frac{c_k}{ik} e^{ikt}.$$

But since

$$f_h(t) = c_0 + \frac{1}{h}\left\{F\left(t+\frac{h}{2}\right) - F\left(t-\frac{h}{2}\right)\right\}$$

we have also

$$f_h(t) = \sum_{-\infty}^{\infty} \frac{2\sin\frac{kh}{2}}{kh} c_k e^{ikt}.$$

Suppose that $f(t) \in L(-\infty, \infty)$ or that $f(t) \in L^2(-\infty, \infty)$. Let

$$\varphi(t) = \mathfrak{F} f(t).$$

$$g(t) = \begin{cases} \frac{1}{h} & \left(-\frac{h}{2} \leqq t \leqq \frac{h}{2}\right), \\ 0 & \left(|t| > \frac{h}{2}\right). \end{cases}$$

Since $f_h(t)$ may be regarded as a folded function derived from $f(t)$ and $g(t)$, and since

$$\mathfrak{F} g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{1}{h} e^{-iut}\, du = \frac{1}{\sqrt{2\pi}} \frac{2\sin\frac{ht}{2}}{ht}$$

it follows, by reason of the theorem on folded functions, that

$$f_h(t) = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \varphi(u) \frac{\sin\frac{hu}{2}}{hu} e^{iut}\, du.$$

This equation holds for nearly every t, if $f(t) \in L(-\infty, \infty)$. It is true everywhere if $f(t) \in L^2(-\infty, \infty)$.

68. Multimonotonic Functions. We begin with a definition. A (real-valued) function $\varphi(x)$ $(a \leqq x < \infty)$ is said to be r-fold monotonic if it can be differentiated r times and it satisfies the relation

$$(-1)^k \varphi^{(k)}(x) \geqq 0 \qquad (a \leqq x < \infty;\ k = 0, 1, 2, \ldots, r) .$$

Ordinarily, in defining multiple monotony, the differentiability of the function is not assumed in advance; and, instead of derivatives, arbitrary finite differences of corresponding orders (with any positive h) are considered. For our future purposes, however, this generalization is of no consequence. We are now going to prove the following theorem, which, strictly speaking, is the combination of two theorems [41]: *If $\varphi(x)$ is bounded in the interval $0 \leqq x < \infty$ and is doubly monotonic, and if $\lim_{x \to \infty} \varphi(x) = 0$, then the integral*

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{\to \infty} \varphi(u) \cos xu \, du$$

exists for every $x \neq 0$. Moreover, the equation

$$\varphi(u) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(x) \cos xu \, du$$

is valid for every positive u.

Proof. Integrating by parts, we find for $x \neq 0$:

$$\int_0^A \varphi(u) \cos xu \, du = \frac{\sin xA}{x} \varphi(A) - \frac{1}{x} \int_0^A \varphi'(u) \sin xu \, du .$$

From our hypothesis, it is evident that $\varphi'(x)$ lies in $L(0, \infty)$. (And, for the same reason, lies also in $L^2(0, \infty)$, a fact of which we shall make use later.) Hence the integral

$$\int_0^A \varphi(u) \cos xu \, du$$

converges to a limit as $A \to \infty$. What is more, the convergence is uniform in every interval which does not contain the origin.

When $x \neq 0$, the existence and continuity of $f(x)$ are thus ensured. Furthermore, the formula

$$-x f(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} \varphi'(u) \sin xu \, du \tag{1}$$

shows that $xf(x) \in L^2(0, \infty)$. Therefore, we have $f(x) \in L(a, \infty)$ for any positive value of a. If we choose $x > 0$, then $f(x)$ can be written in the form:

$$f(x) = \sqrt{\frac{2}{\pi}} \sum_{n=0}^{\infty} \int_0^{\frac{\pi}{2}} \left\{ \varphi\left(\frac{v + 2n\pi}{x}\right) - \varphi\left(\frac{\pi - v + 2n}{x}\right) - \right.$$

$$\left. - \varphi\left(\frac{\pi + v + 2n\pi}{x}\right) + \varphi\left(\frac{2\pi - v + 2n\pi}{x}\right) \right\} \frac{\cos v}{x}\, dv.$$

Now, the expression in the braces is not negative because $\varphi(x)$ is doubly monotonic. Consequently, $f(x) \geq 0$.

By applying Plancherel's theorem to (1), we get

$$\varphi(0) - \varphi(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} (1 - \cos ux) f(u)\, du. \tag{2}$$

Whence, for an arbitrary $\varepsilon > 0$, it follows that

$$\varphi(0) - \varphi(x) \geqq \sqrt{\frac{2}{\pi}} \int_{\varepsilon}^{\infty} (1 - \cos ux) f(u)\, du.$$

If ε is held fixed and x is allowed to become infinite, we obtain, by virtue of the Riemann-Lebesgue theorem,

$$\varphi(0) \geqq \sqrt{\frac{2}{\pi}} \int_{\varepsilon}^{\infty} f(u)\, du.$$

If we now allow ε to approach zero, we find that $f(x) \in L(0, \infty)$. Hence — again by virtue of (2) and the Riemann-Lebesgue theorem — we have

$$\varphi(0) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(u)\, du \text{ und } \varphi(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(u) \cos ux\, du.$$

This proves the theorem.

69. Conjugate Functions. Frequently, we find it necessary to consider, along with a trigonometric series

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt) \tag{1}$$

also its so-called conjugate [38]

$$\frac{b_0}{2} + \sum_{k=1}^{\infty} (b_k \cos kt - a_k \sin kt) \,. \tag{2}$$

Here the constant b_0 remains wholly arbitrary. Suppose, now, that (1) is the Fourier series of a function $f(t)$. In the case where (2) is likewise the Fourier series of a function, the latter function is said to be a conjugate to $f(t)$ [8] and is denoted by $\tilde{f}(t)$. From the discussion in Section 55, it follows that we have as an example

$$\psi_1(t) \equiv \ln\left(2\cos\frac{t}{2}\right) = \tilde{\varphi}_1(t)$$

with

$$\varphi_1(t) = \frac{t}{2} \qquad (-\pi < t < \pi) \,.$$

A quite simple case is at hand if the condition

$$\sum_{k=1}^{\infty} \{|a_k|^2 + |b_k|^2\} < \infty$$

is fulfilled. In this case, (1) and (2) are the Fourier series of functions in $L^2(0, 2\pi)$. If, now, $b_0 = a_0$, then we have

$$\|\tilde{f}\| = \|f\| ,$$

that is to say we obtain a unitary operator which transforms $f(t)$ into $\tilde{f}(t)$.

When passing from functions on finite intervals to functions along the entire real axis, we assume first that $f(t) \in L^2(-\infty, \infty)$. Suppose now that $\varphi(t)$ is the Fourier transform of $f(t)$:

$$f(t) = \underset{\substack{a\to-\infty \\ b\to\infty}}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_a^b \varphi(u)\, e^{itu}\, du \,. \tag{3}$$

The function [41]

$$\tilde{f}(t) = \underset{\substack{a\to-\infty \\ b\to\infty}}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_a^b \varphi(u)\, (i \operatorname{sign} u)\, e^{itu}\, du \tag{4}$$

is now called the conjugate of $f(t)$. It is clear that

$$\|\tilde{f}\| = \|f\| \,.$$

This definition is entirely analogous to the foregoing definition for series, since, from

$$\int_{-n}^{n} e^{itu}\, \varphi(u)\, du = \int_0^n \{A(u) \cos tu + B(u) \sin tu\}\, du$$

[8] At times we speak also of the trigonometric conjugate to stress the fact that the complex conjugate function of $f(t)$ is not intended at all.

we obtain of course

$$\int_{-n}^{n} e^{itu}\,(i\,\mathrm{sign}\,u)\,\varphi(u)\,du = \int_{0}^{n} \{B(u)\cos tu - A(u)\sin tu\}\,du.$$

By means of an appropriate generalization, the concept of conjugate function can be extended also to other classes of functions. In the future we will restrict ourselves to the class W^2 of all those measurable functions $f(t)$ $(-\infty < t < \infty)$ for which [52] the relation

$$\frac{f(t)}{1+|t|} \in L^2(-\infty, \infty)$$

is satisfied. Suppose, then, that $f(t) \in W^2$. Let us now introduce $\psi(t)$ as the Fourier transform of the function

$$\frac{f(t)}{t-i};$$

then we have

$$f(t) = (t-i)\,\frac{1}{\sqrt{2\pi}}\,\underset{\substack{a\to-\infty\\ b\to\infty}}{\text{l. i. m.}} \int_{a}^{b} e^{itu}\,\psi(u)\,du. \tag{5}$$

By definition, we have further

$$\tilde{f}(t) = (t-i)\,\frac{1}{\sqrt{2\pi}}\,\underset{\substack{a\to-\infty\\ b\to\infty}}{\text{l. i. m.}} \int_{a}^{b} e^{itu}\,(i\,\mathrm{sign}\,u)\,\psi(u)\,du. \tag{6}$$

Here again we meet a certain operator which is unitary in the sense that

$$\int_{-\infty}^{\infty} \frac{|\tilde{f}(t)|^2}{1+t^2}\,dt = \int_{-\infty}^{\infty} \frac{|f(t)|^2}{1+t^2}\,dt\,.$$

Every periodic function $f(t)$ belonging to $L^2(0, 2\pi)$ lies also in W^2. Likewise, every function belonging to $L^2(-\infty, \infty)$ is also contained in W^2. It is therefore necessary to verify that the conjugate function in the general sense, defined by the formulas (5), (6), is essentially the same as the conjugate function in the first sense (in other words, that these two coincide almost everywhere save for an additive constant) both in the case where $f(t)$ is periodic and belongs to $L^2(0, 2\pi)$ and in the case where $f(t) \in L^2(-\infty, \infty)$.

To that end, let us observe first of all:

$$\mathfrak{F}\left\{\frac{1}{t-i}\right\} = g(t) = \begin{cases} i\sqrt{2\pi}\,e^{t} & (t<0),\\ 0 & (t \geqq 0). \end{cases}$$

Suppose, now, that $f(t) \in L^2(-\infty, \infty)$. In conformity with the formulas (3), (5), we will now introduce the functions $\varphi(t)$, $\psi(t)$, of which the second is obviously continuous along the entire axis. Furthermore, we will denote the function conjugate to $f(t)$ and defined in the formula (4) by $\tilde{f}_0(t)$, and the function conjugate to $f(t)$ and defined in the formula (6) by $\tilde{f}(t)$.

By reason of the theorem on the Faltung of functions, we have

$$\psi(t) \equiv \mathfrak{F}\left\{f(t) \cdot \frac{1}{t-i}\right\} = i\int_{-\infty}^{0} \varphi(t-v)\, e^{v}\, dv.$$

Similarly,

$$\omega(t) \equiv \mathfrak{F}\left\{\tilde{f}_0(t) \cdot \frac{1}{t-i}\right\} = i\int_{-\infty}^{0} \varphi(t-v)\,[i \operatorname{sign}(t-v)]\, e^{v}\, dv.$$

Hence if t is not negative, we obtain

$$\omega(t) = i\psi(t).$$

If t is negative, however, we obtain

$$\omega(t) = i\int_{-\infty}^{t} \varphi(t-v)\, ie^{v}\, dv - i\int_{t}^{0} \varphi(t-v)\, ie^{v}\, dv =$$

$$= 2i\int_{-\infty}^{t} i\varphi(t-v)\, e^{v}\, dv - i\int_{-\infty}^{0} i\varphi(t-v)\, e^{v}\, dv = 2i\psi(0)\, e^{t} - i\psi(t).$$

Now, since

$$\tilde{f}_0(t) = (t-i)\frac{1}{\sqrt{2\pi}}\,\text{l. i. m.}_{\substack{a\to-\infty\\ b\to\infty}} \int_{a}^{b} \omega(u)\, e^{itu}\, du$$

it follows that

$$\tilde{f}_0(t) = (t-i)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} 2i\psi(0)\, e^{u}e^{itu}\, du +$$

$$+ (t-i)\frac{1}{\sqrt{2\pi}}\,\text{l. i. m.}_{\substack{a\to-\infty\\ b\to\infty}} \int_{a}^{b} e^{itu}\,(i \operatorname{sign} u)\,\psi(u)\, du$$

or

$$\tilde{f}_0(t) = \frac{2\psi(0)}{\sqrt{2\pi}} + \tilde{f}(t).$$

This proves our proposition. The case where $f(t)$ is periodic and lies in $L^2(0, 2\pi)$ is treated in a similar fashion.

In conclusion, it should be noted that it is possible to begin at the outset with $\frac{f(t)-C}{t-\gamma}$ instead of $\frac{f(t)}{t-i}$, where C and γ denote any constants for which the fraction $\frac{f(t)-C}{t-\gamma}$ belongs to $L^2(-\infty, \infty)$. If this substitution were made, the conjugate function would be the same as above except for an additive constant.

CHAPTER IV

SOME EXTREMAL PROPERTIES OF INTEGRAL TRANSCENDENTAL FUNCTIONS OF THE EXPONENTIAL TYPE

70. Integral Functions of the Exponential Type. We say that an integral transcendental function $f(z)$ is of the exponential type if the relation

$$\overline{\lim_{r\to\infty}} \frac{\ln M(r)}{r} < \infty$$

is valid, where

$$M(r) = \max_{|z|=r} |f(z)| .$$

The number

$$\sigma = \overline{\lim_{r\to\infty}} \frac{\ln M(r)}{r}$$

is called the exponent of $f(z)$.

On the basis of the definition just given, $f(z)$ is an integral transcendental function of the exponential type and with the exponent σ (or more briefly: a function of the class $\mathbf{E}_\sigma$) if, corresponding to every $\varepsilon > 0$, the inequality

$$|f(z)| < e^{(\sigma+\varepsilon)|z|}$$

is satisfied for every sufficiently large $|z|$ and if, furthermore, the inequality

$$|f(z)| > e^{(\sigma-\varepsilon)|z|}$$

holds for certain arbitrarily large values of $|z|$.

Let us present a few functions of the class $\mathbf{E}_\sigma$:

$$P(z)\, e^{\sigma z},\ \frac{e^{\sigma z}-1}{z} P(z),\qquad P(z)\cos(\sigma z+\alpha);$$

here $P(z)$ denotes an arbitrary polynomial in z.

Now, let

$$f(z) = c_0 + c_1 z + c_2 z^2 + \dots \tag{1}$$

be a given integral transcendental function. We are going to show that each of the two inequalities

(2) $$\lim_{r\to\infty} \frac{\ln M(r)}{r} \leqq \sigma,$$

(3) $$\lim_{n\to\infty} \frac{n}{e} \sqrt[n]{|c_n|} = \overline{\lim_{n\to\infty}} \sqrt[n]{n!\,|c_n|} \leqq \sigma$$

is a consequence of the other.

Whence it will follow that $f(z)$ belongs to the class $\mathbf{E}_\sigma$ if and only if the relation

$$\lim_{n\to\infty} \sqrt[n]{n!\,|c_n|} = \sigma$$

is valid.

Now, in order to prove that (2) follows from (3), we will rely upon Cauchy's inequality

$$|c_n| \leqq \frac{M(r)}{r^n} \qquad (n = 0, 1, 2, \ldots)$$

and set $r = \dfrac{n}{\sigma + \varepsilon}$. Here ε is an arbitrarily chosen positive number. Then, according to (2), the inequality

$$|c_n| \leqq \frac{e^n (\sigma + s)^n}{n^n}$$

is satisfied for every $n \geq N(\varepsilon)$; this inequality can be put in the form

$$\frac{n}{e} \sqrt[n]{|c_n|} \leqq \sigma + \varepsilon$$

from which (3) follows immediately.

Let us now prove that, conversely, (2) is a consequence of (3). For that purpose, we will start with the inequality

$$M(r) \leqq |c_0| + |c_1| r + |c_2| r^2 + \ldots$$

By virtue of (3), we have for every $n \geq N(\varepsilon)$ the estimate

$$|c_n| \leqq \frac{\left(\sigma + \dfrac{\varepsilon}{2}\right)^n}{n!}$$

and consequently also

$$M(r) < \sum_{k=0}^{N(\varepsilon)-1} |c_k| r^k + e^{r\left(\sigma + \frac{\varepsilon}{2}\right)}.$$

As a result thereof, we have the inequality

$$M(r) < e^{r(\sigma+\varepsilon)},$$

for all values of $r \geq R(\varepsilon)$; this means, however, that

$$\frac{\ln M(r)}{r} < \sigma + \varepsilon .$$

Thus (2) has been derived from (3).

From the theorem just proven, it follows, among other things, that the derivative of a function in $\mathbf{E}_\sigma$ belongs again to $\mathbf{E}_\sigma$.

71. The Borel Transformation [53]. Let

$$f(z) = a_0 + \frac{a_1}{1!} z + \frac{a_2}{2!} z^2 + \dots \tag{1}$$

be a given function of the class $\mathbf{E}_\sigma (\sigma > 0)$. Then we have

$$\overline{\lim_{n \to \infty}} \sqrt[n]{|a_n|} = \sigma .$$

This relation says, however, that the power series

$$\frac{a_0}{z} + \frac{a_1}{z^2} + \frac{a_2}{z^3} + \dots = g(z) \tag{2}$$

has σ as the radius of divergence. [1] Thus the functions of class $\mathbf{E}_\sigma$ and the power series of the form (2) with σ as the radius of divergence have been brought into a one-to-one correspondence.

Furthermore, let us consider the integral

$$\int_0^\infty f(\zeta)\, e^{-z\zeta}\, d\zeta = g(z; \Theta) \tag{3}$$

extended over the half-line $\zeta = \varrho e^{-i\Theta} (0 \leq \varrho < \infty)$, which forms the angle $-\Theta$ with the real axis. That half-plane which does not contain the point $z = 0$ and whose boundary is tangent to the circle $|z| = \sigma$ at the point $\sigma e^{i\Theta}$ will be denoted by Δ_Θ.

Let us now show that the integral (3) converges absolutely and uniformly in every closed bounded subset of the open half-plane Δ_Θ and hence represents a regular analytic function on Δ_Θ.

Suppose that the closed region D lies in the interior of the half-plane Δ_Θ. Then the distance of every point z in D, from the boundary of the half-plane Δ_Θ, is greater than a fixed positive δ. In other words, for all points $z = x + iy$ of D we have

$$\Re(ze^{-i\Theta}) = x \cos\Theta + y \sin\Theta > \sigma + \delta .$$

[1] Translator's Note: We designate the radius of convergence of $g\left(\frac{1}{z}\right)$ as the radius of divergence of $g(z)$.

On the other hand, however, since $f(z)$ belongs to class $\mathbf{E}_\sigma$, there exists a constant C such that for every z we have the estimate

$$|f(z)| \leqq Ce^{\left(\sigma+\frac{\delta}{2}\right)z}$$

Therefore, the inequality

$$|f(\zeta)\, e^{-z\zeta}| \leqq Ce^{\left(\sigma+\frac{\delta}{2}\right)\varrho} e^{-\varrho(\sigma+\delta)} = Ce^{-\frac{\delta}{2}\varrho}$$

is satisfied on the half-line $\zeta = \varrho e^{-i\Theta}(0 \leq \varrho < \infty)$ by the points z which are situated in D. If we take notice, moreover, that

$$\int_0^\infty Ce^{-\frac{\delta}{2}\varrho}\, d\varrho = \frac{2C}{\delta} < \infty,$$

our proof is complete.

The integral (3) is called the Borel transform of the function $f(z)$ for the half-plane Δ_Θ.

It is possible that the whole region of regularity of the function $g(z;\Theta)$ does not coincide with the half-plane Δ_Θ, but contains the latter as a proper subset. This fact is so important because, as we shall soon prove, the equation

$$g(z;\Theta) = g(z) \tag{4}$$

holds at every point z in Δ_Θ. $g(z)$ in (4) is the same function that appears above in (2). Hence in certain cases the Borel transformation can achieve the analytic extension of the function $g(z)$, which originally was defined only in the region $|z| > \sigma$.

We are now going to prove that every point z in Δ_Θ satisfies the equation (4). Let (cf. above)

$$\Re(e^{-i\Theta}z) > \sigma + \delta \qquad (\delta > 0)\,. \tag{5}$$

It suffices to prove the absolute and uniform convergence of the series

$$f(\zeta)\, e^{-z\zeta\frac{\sigma+\frac{\delta}{2}}{\sigma+\delta}} = \sum_{k=0}^{\infty} \frac{a_k}{k!}\zeta^k e^{-z\zeta\frac{\sigma+\frac{\delta}{2}}{\sigma+\delta}}$$

along the ray $\zeta = \varrho e^{-i\Theta}(0 \leq \varrho < \infty)$. Then the equation (4) could be obtained by termwise integration. Since $f(z)$ belongs to the class $\mathbf{E}_\sigma$, we can conclude by reason of (3), Section 70, that there exists a constant C which satisfies the inequality

$$|a_k| \leqq C\left(\sigma + \frac{\delta}{4}\right)^k$$

for every value of k. Consequently, we have at all points of the straight half-line $\zeta = \varrho e^{-i\Theta} (0 \leq \varrho < \infty)$ the estimate

$$\left| \frac{a_k}{k!} \zeta^k e^{-z\zeta \frac{\sigma+\frac{\delta}{2}}{\sigma+\delta}} \right| \leqq C \frac{\left(\sigma + \frac{\delta}{4}\right)^k}{k!} \varrho^k e^{-\left(\sigma+\frac{\delta}{2}\right)\varrho} \leqq C \left(\frac{\sigma + \frac{\delta}{4}}{\sigma + \frac{\delta}{2}} \right)^k \quad (k = 0, 1, 2, \ldots),$$

provided (5) is satisfied. This proves the proposition.

72. The Theorem of Wiener and Paley. We will denote by W_σ the totality of all those integral transcendental functions $f(z)$ of the exponential type which have an exponent $\leq \sigma$ and satisfy the inequality

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx < \infty \, .$$

Wiener and Paley [54] proved that the class W_σ coincides with the set of those functions $f(z)$ which admit a representation of the form

$$f(z) = \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} e^{iuz} \varphi(u) \, du \tag{1}$$

with $\varphi(u) \in L^2(-\sigma, \sigma)$.

Suppose, now, that $\varphi(u) \in L^2(-\sigma, \sigma)$. It is very easy to show that every function of the form (1) belongs to the class W_σ. In fact: First of all, it follows from (1) that

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \int_{-\sigma}^{\sigma} |\varphi(u)|^2 \, du < \infty.$$

If we set $z = x + iy$, we obtain

$$f(z) = f(x + iy) = \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} e^{iuz} e^{-uy} \varphi(u) \, du.$$

With the aid of Schwarz's inequality, we can conclude from the previous equation that

$$|f(z)| \leqq \frac{1}{\sqrt{2\pi}} \sqrt{\int_{-\sigma}^{\sigma} |\varphi(u)|^2 \, du} \sqrt{\int_{-\sigma}^{\sigma} e^{-2uy} \, du} =$$

$$= C_1 \sqrt{\frac{e^{2\sigma y} - e^{-2\sigma y}}{y}} \leqq C_2 e^{\sigma |y|} \leqq C_2 e^{\sigma |z|} \, .$$

That, conversely, every function in W_σ admits a representation of the form (1), is more difficult to prove. Plancherel and Pólya [55] have given an elegant proof thereof, which we will now repeat:

The problem consists in proving the following statement: "From the condition $f(z) \in W_\sigma$ it follows that the Fourier transform $\varphi(x)$ of the function $f(x)$ assumes the value zero for almost every $x > \sigma$ and almost every $x < -\sigma$." If this is in fact the case, then

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-0}^{0} e^{iux}\, \varphi(u)\, du.$$

But then the integral function

$$\frac{1}{\sqrt{2\pi}} \int_{-0}^{0} e^{iuz}\, \varphi(u)\, du$$

will coincide with $f(z)$ for $z = x$ $(-\infty < x < \infty)$ and hence will be identical with $f(z)$.

Let us now go back to the function

$$g(z) = \frac{a_0}{z} + \frac{a_1}{z^2} + \frac{a_2}{z^3} + \dots \qquad (|z| > \sigma)$$

which is associated with $f(z)$. We will prove with the aid of the Borel transform of $f(z)$ that $g(z)$ is regular in the complex plane which is cut along the imaginary axis from $-i\sigma$ to $+i\sigma$. For that purpose, let us construct the Borel transform for $\Theta = 0, \pi$. Thus we obtain the expressions

$$g(x+iy) = \int_{0}^{\infty} f(\xi)\, e^{-\xi(x+iy)}\, d\xi \equiv g_+(x+iy), \tag{2}$$

$$g(x+iy) = -\int_{0}^{\infty} f(-\xi)\, e^{\xi(x+iy)}\, d\xi = \tag{3}$$

$$= -\int_{-\infty}^{0} f(\xi)\, e^{-\xi(x+iy)}\, d\xi \equiv g_-(x+iy),$$

which, by reason of the results in the preceding section, are valid for either $x > \sigma$ or $x < -\sigma$, as the case may be. If we observe further that $f(z)$ lies in $L^2(-\infty, \infty)$, we can conclude at once, with the aid of Schwarz's inequality, that both integrals converge uniformly for $x \geqq \varepsilon > 0$, $x \leqq \varepsilon < 0$, respectively. That, however, means: Here Borel's transform actually represents an analytic extension of the function $g(z)$, which had been defined by the above series only for $|z| > \sigma$.

Thus $g(z)$ is regular (and hence single-valued) on the whole plane which is cut along the imaginary axis from $-i\sigma$ to $+i\sigma$.

Whence it follows that for $y > \sigma$ and $y < -\sigma$ the following limits exist and are equal to $g(iy)$:

$$g_+(iy+0) = \lim_{\varepsilon \to 0} g_+(iy+\varepsilon), \qquad (\varepsilon > 0)$$

$$g_-(iy-0) = \lim_{\varepsilon \to 0} g_-(iy-\varepsilon).$$

By virtue of (2), $g_+(iy+\varepsilon)$ coincides with the Fourier transform of the function

$$f_+(x;\varepsilon) = \begin{cases} f(x)\, e^{-\varepsilon x} & (x \geqq 0), \\ 0 & (x < 0) \end{cases}$$

except for the factor $\frac{1}{\sqrt{2\pi}}$. This function lies in L^2. Therefore the expression

$$\int_{-\infty}^{\infty} |f_+(x;\varepsilon) - f_+(x;0)|^2\, dx = \int_0^{\infty} |f(x)|^2 (1 - e^{-\varepsilon x})^2\, dx$$

converges to zero if ε approaches zero. Consequently, we have

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{2\pi}} g_+(iy+\varepsilon) - \varphi_+(y) \right|^2 dy = 0$$

with

$$\varphi_+(y) = \text{l.i.m.}_{n \to \infty} \frac{1}{\sqrt{2\pi}} \int_0^n f(x)\, e^{-ixy}\, dx.$$

But since, as we have shown above, the function $g_+(iy+\varepsilon)$ has an *ordinary* limit for $y > \sigma$ and $y < -\sigma$, if ε approaches 0, then, for almost every $y > \sigma$ and almost every $y < -\sigma$, we have the equation

$$\varphi_+(y) = \frac{1}{\sqrt{2\pi}} g_+(iy+0) = \frac{1}{\sqrt{2\pi}} g(iy).$$

In an analogous manner, it can be shown that (in the same sense)

$$\varphi_-(y) = -\frac{1}{\sqrt{2\pi}} g_-(iy-0) = -\frac{1}{\sqrt{2\pi}} g(iy)$$

where this time we set

$$\varphi_-(y) = \text{l.i.m.}_{n \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-n}^{0} f(x)\, e^{-ixy}\, dx\,.$$

Hence the relation

$$\varphi(y) = \underset{\substack{a\to-\infty\\ b\to\infty}}{\text{l. i. m.}} \frac{1}{\sqrt{2\pi}} \int_a^b f(x)\, e^{-ixy}\, dx = \varphi_+(y) + \varphi_-(y) =$$

$$= \frac{1}{\sqrt{2\pi}} \{g(iy) - g(iy)\} = 0$$

is valid for almost every $y > \sigma$ and almost every $y < -\sigma$. This proves the theorem.

73. Integral Functions of the Exponential Type which are Bounded along the Real Axis. We will denote by B_σ the totality of those integral transcendental functions of the exponential type whose exponent is not greater than σ and which satisfy the inequality

$$\sup_{-\infty<x<\infty} |f(x)| < \infty .$$

As an example of a function belonging to B_σ, let us employ the integral

$$\int_{-\sigma}^{\sigma} e^{iuz}\, d\omega(u), \tag{1}$$

where any function of bounded variation may be substituted for $\omega(u)$. In particular, the sums

$$\sum_k c_k e^{i\alpha_k z} \qquad (-\sigma \leqq \alpha_k \leqq \sigma),$$

such that

$$\sum_k |c_k| < \infty,$$

and likewise the integrals

$$\int_{-\sigma}^{\sigma} e^{iuz}\, c(u)\, du, \tag{2}$$

which satisfy the inequality

$$\int_{-\sigma}^{\sigma} |c(u)|\, du < \infty,$$

accordingly belong to B_σ and consequently so do all the functions in W_σ.

S.N. Bernstein [56] was the first to call attention to the class B_σ. To him is due a remarkable theorem about this class; we will formulate this theorem

presently. It plays an important role in problems of the approximation of *continuous functions.*

If $f(z)$ belongs to B_σ, then the function

$$\frac{f(z)-f(0)}{z}$$

obviously lies in W_σ. Hence it follows from the theorem of Wiener and Paley that every function $f(z)$ in B_σ admits a representation of the form

$$f(z)=f(0)+\frac{z}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma} e^{iuz}\psi(u)\,du$$

with $\psi(u)\in L^2(-\sigma,\sigma)$. This representation, in conjunction with the considerations of Section 69, permits us to associate with every function $f(z)$ in B_σ a conjugate function [2]

$$\tilde{f}(z)=\frac{z}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma}(i\,\operatorname{sign} u)\,e^{iuz}\psi(u)\,du$$

which is likewise an integral transcendental function with an exponent $\leq\sigma$. However, we cannot assert that $\tilde{f}(z)$ also belongs to the class B_σ.

In what follows we will frequently consider, together with a function $f(z)$ lying in B_σ, also its conjugate function $\tilde{f}(z)$.

The above-mentioned theorem of S. N. Bernstein says: *If the function $f(z)$ lies in B_σ, then $f(z)$ satisfies the inequality*

$$\sup_{-\infty<x<\infty}|f'(x)|\leqq\sigma\sup_{-\infty<x<\infty}|f(x)|.$$

The equality sign is admitted, but only in the case of functions of the form

$$f_0(z)=ae^{i\sigma z}+be^{-i\sigma z},$$

where a, b are constants.

It follows from the aforementioned theorem, often called Bernstein's inequality, that all the functions $f(z)$ belonging to B_σ satisfy the inequality

$$\sup_{-\infty<x<\infty}|f^{(k)}(x)|\leqq\sigma^k\sup_{-\infty<x<\infty}|f(x)|,$$

where k is any positive integer.

---

[2] The indeterminateness in the definition of the conjugate function is removed here through the normalization

$$\tilde{f}(0)=0.$$

It can be seen without difficulty that the following converse is true: *If* $f(x)$ $(-\infty < x < +\infty)$ *can be differentiated as often as we please and if it satisfies the inequality*

$$\sup_{-\infty<x<\infty} |f^{(k)}(x)| \leqq \sigma^k M \qquad (k = 0, 1, 2, \ldots)$$

then there exists a function $f(z)$ *in* B_σ *such that* $f(z)$ *coincides with* $f(x)$ *when* $z = x$ $(-\infty < x < \infty)$.

In fact: The Maclaurin series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} z^k$$

converges for all finite values of z and represents the function $f(x)$ on the real axis. The absolute value of the series, however, does not exceed

$$M \sum_{k=0}^{\infty} \frac{\sigma^k}{k!} |z|^k = Me^{\sigma|z|}.$$

Even still more can be proven, namely: The expression

$$|f(z)| \leqq Me^{\sigma|y|} \tag{3}$$

is valid for any value of $z = x + iy$, since

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!} (iy)^k,$$

which means, however, that

$$|f(z)| \leqq M \sum_{k=0}^{\infty} \frac{\sigma^k}{k!} |y|^k = Me^{\sigma|y|}.$$

In a similar manner, we can prove the inequality

$$|f'(z)| \leqq M \sigma e^{\sigma|y|}.$$

These two inequalities now show: Suppose that $\mathfrak{M} \subset B_\sigma$, and, besides, that all the functions in $\mathfrak{M}$ are uniformly bounded on the real axis. Then the functions in $\mathfrak{M}$ are equicontinuous in every bounded point set of the complex plane. Hence, according to a well-known theorem in analysis, every infinite sequence of functions in $\mathfrak{M}$ contains a partial sequence which converges uniformly in every bounded subset of the complex plane. This property is known as "compactness" and will be utilized subsequently. It will be noted further that the limit function satisfies the inequality (3) everywhere, and therefore is contained in B_σ.

S. N. Bernstein proved the inequality named after him, first [57] for trigonometric sums, and then later [56] for arbitrary functions in B_σ.

We owe to G. Szegö [58] an important generalization of Bernstein's inequality for trigonometric sums, which is tied in with the following problem.

Let $\mathfrak{M}_n$ denote the totality of all the trigonometric sums

$$f(z) = \sum_{k=-n}^{n} c_k e^{ikz}$$

which satisfy the condition

$$\max_{0 \leqq x \leqq 2\pi} |f(x)| \leqq 1 .$$

Furthermore, suppose given a sequence λ_k $(k = 0, \pm 1, \pm 2, \ldots, \pm n)$ of real or complex numbers. With the aid of this sequence, we can associate with every trigonometric sum $f(z)$ belonging to $\mathfrak{M}_n$ a trigonometric sum

$$f_\lambda(z) = \sum_{k=-n}^{n} \lambda_k c_k e^{ikz}$$

(the λ-transform of the trigonometric sum $f(z)$). Let now the value of

$$\sup_{f \in \mathfrak{M}_n} \max_{0 \leqq 0 \leqq 2\pi} |f_\lambda(x)| = \mu_\lambda$$

be given. Then those trigonometric sums are to be determined (in case such exist at all) for which this supremum was assumed.

S. N. Bernstein's problem is a special case of this general "multiplication problem," and the theorem of Bernstein asserts that the quantity μ_λ is equal to n for $\lambda_k = ik$ $(k = 0, \pm 1, \pm 2, \ldots, \pm n)$ and that the trigonometric polynomials of the form

$$ae^{inz} + be^{-inz} \qquad (|a| + |b| = 1)$$

are extremal functions.

The multiplication problem can be presented analogously in the continuous case as well. Then, instead of trigonometric sums, we shall have trigonometric integrals of the form (1) or (2).

M. G. Krein had supplemented and completed the investigations of G. Szegö. Some of Krein's results will be given later (cf. Section 77).

74. S. N. Bernstein's Inequality. We are now going to prove two theorems. Each of them contains Bernstein's inequality as a special case.

THEOREM 1. *If $f(z)$ belongs to B_σ, then the inequality*

$$\sup_{-\infty < x < \infty} |\sin \alpha \cdot f'(x) - \sigma \cos \alpha \cdot f(x)| \leqq \sigma \sup_{-\infty < x < \infty} |f(x)|$$

is satisfied for every real value of α.

The equality sign is valid if and only if $f(z)$ has the form

$$f(z) = ae^{i\sigma z} + be^{-i\sigma z} .$$

THEOREM 2. *If $f(z)$ belongs to B_σ, then the inequality*

$$\sup_{-\infty<x<\infty} |\sin\alpha\cdot f'(x)+\cos\alpha f'(x)| \leqq \sigma \sup_{-\infty<x<\infty} |f(x)|$$

is satisfied for every real value of α.
The equality sign holds for the same functions as in Theorem 1.

Proof: Let

$$(1)\qquad f(z)=f(0)+\frac{z}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma} e^{iuz}\psi(u)\,du = f(0)+\frac{1}{i\sqrt{2\pi}}\int_{-\sigma}^{\sigma}\psi(u)\frac{\partial}{\partial u}e^{iuz}\,du$$

and accordingly

$$(1')\qquad \tilde{f}(z)=\frac{z}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma} i\,\operatorname{sign} u\cdot e^{iuz}\psi(u)\,du,$$

where $\psi(u)\in L^2$. What interests us are the expressions

$$U_1(x)=\sin\alpha\cdot f'(x)-\sigma\cos\alpha\cdot f(x),$$
$$U_2(x)=\sin\alpha\cdot f'(x)+\cos\alpha\cdot\tilde{f}'(x).$$

These can now be put into the form

$$U_1(x)=-\sigma\cos\alpha\cdot f(0)+\frac{1}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma}\psi(u)\frac{\partial}{\partial u}\{e^{iux}(u\sin\alpha+i\sigma\cos\alpha)\}\,du,$$

$$U_2(x)=\frac{1}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma}\psi(u)\frac{\partial}{\partial u}\{e^{iux}(u\sin\alpha+i|u|\cos\alpha)\}\,du.$$

Let us now develop the functions

$$V_1(u)=-ie^{\frac{i\alpha u}{\sigma}}(u\sin\alpha+i\sigma\cos\alpha), \qquad (-\sigma\leqq u\leqq\sigma)$$

$$V_2(u)=-ie^{\frac{i\alpha u}{\sigma}}(u\sin\alpha+i|u|\cos\alpha)$$

in Fourier series. These series then have the following form:

$$V_1(u)=\sigma\sin^2\alpha\sum_{k=-\infty}^{\infty}(-1)^k\frac{1}{(\alpha-k\pi)^2}e^{\frac{ik\pi}{\sigma}u},$$

$$V_2(u)=2\sigma\sum_{k=-\infty}^{\infty}(-1)^k\frac{\sin^2\frac{\alpha-k\pi}{2}}{(\alpha-k\pi)^2}e^{\frac{ik\pi}{\sigma}u}.$$

Accordingly we obtain with the aid of (1) the following expressions for the functions $U_1(x)$, $U_2(x)$:

$$U_1(x) = -\sigma\cos\alpha\cdot f(0) +$$

$$+\frac{i\sigma\sin^2\alpha}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma}\psi(u)\frac{\partial}{\partial u}\left\{\sum_{k=-\infty}^{\infty}\frac{(-1)^k}{(\alpha-k\pi)^2}e^{iu\left(\frac{k\pi-\alpha}{\sigma}+x\right)}\right\}du =$$

$$= -\sigma\cos\alpha\cdot f(0) - \sigma\sin^2\alpha\sum_{k=-\infty}^{\infty}\frac{(-1)^k}{(\alpha-k\pi)^2}\left\{f\left(\frac{k\pi-\alpha}{\sigma}+x\right)-f(0)\right\} =$$

$$= \sigma\sin^2\alpha\sum_{k=-\infty}^{\infty}\frac{(-1)^{k-1}}{(\alpha-k\pi)^2}f\left(\frac{k\pi-\alpha}{\sigma}+x\right),$$

since

$$\sum_{k=-\infty}^{\infty}\frac{(-1)^k}{(\alpha-k\pi)^2} = -\frac{d}{d\alpha}\frac{1}{\sin\alpha} = \frac{\cos\alpha}{\sin^2\alpha}$$

and

$$U_2(x) = \frac{2i\sigma}{\sqrt{2\pi}}\int_{-\sigma}^{\sigma}\psi(u)\frac{\partial}{\partial u}\left\{\sum_{k=-\infty}^{\infty}(-1)^k\frac{\sin^2\frac{\alpha-k\pi}{2}}{(\alpha-k\pi)^2}e^{iu\left(\frac{k\pi-\alpha}{\sigma}+x\right)}\right\}du =$$

$$= -2\sigma\sum_{k=-\infty}^{\infty}(-1)^k\frac{\sin^2\frac{\alpha-k\pi}{2}}{(\alpha-k\pi)^2}\left\{f\left(\frac{k\pi-\alpha}{\sigma}+x\right)-f(0)\right\} =$$

$$= 2\sigma\sum_{k=-\infty}^{\infty}(-1)^{k-1}\frac{\sin^2\frac{\alpha-k\pi}{2}}{(\alpha-k\pi)^2}f\left(\frac{k\pi-\alpha}{\sigma}+x\right),$$

because

$$\sum_{k=-\infty}^{\infty}(-1)^k\frac{\sin^2\frac{\alpha-k\pi}{2}}{(\alpha-k\pi)^2} = 0.$$

Thus we have found the following "interpolation formulas" [3] for the functions in B_σ:

$$\text{(I)}\qquad \sin\alpha\cdot f'(x) - \sigma\cos\alpha\cdot f(x) = \sigma\sum_{k=-\infty}^{\infty}(-1)^{k-1}\frac{\sin^2\alpha}{(\alpha-k\pi)^2}f\left(\frac{k\pi-\alpha}{\sigma}+x\right),$$

[3] The first of these formulas has been known for a long time (cf. e.g. [8], vol. 1, p. 143). Now there is no further difficulty in completing the proof.

(II) $\sin\alpha \cdot f'(x) + \cos\alpha \cdot \tilde{f}'(x) =$

$$= \sigma \sum_{k=-\infty}^{\infty} (-1)^{k-1} \frac{2\sin^2 \frac{\alpha - k\pi}{2}}{(\alpha - k\pi)^2} f\left(\frac{k\pi - \alpha}{\sigma} + x\right).$$

Let

$$\sup_{-\infty < x < \infty} |f(x)| = M.$$

Then, by virtue of (I), we have

$$\sup_{-\infty < x < \infty} |\sin\alpha \cdot f'(x) - \sigma\cos\alpha \cdot f(x)| \leqq \sigma M \sum_{k=-\infty}^{\infty} \frac{\sin^2\alpha}{(\alpha - k\pi)^2} = \sigma M,$$

and the equality sign is taken in the case where, for any value of x — which we will denote by x_0 — we have the relations

$$f\left(\frac{k\pi - \alpha}{\sigma} + x_0\right) = (-1)^k M e^{i\beta} \qquad (k = 0, \pm 1, \pm 2, \ldots)$$

with a real constant β. If, however, these equations are satisfied, then (I) can be brought into the form

$$-\sin\sigma\left(x - x_0 + \frac{\alpha}{\sigma}\right) f'(x) + \sigma\cos\sigma\left(x - x_0 + \frac{\alpha}{\sigma}\right) f(x) =$$

$$= \sigma M e^{i\beta} \sum_{k=-\infty}^{\infty} \frac{\sin^2\sigma\left(x - x_0 + \frac{\alpha}{\sigma}\right)}{\left[\sigma\left(x - x_0 + \frac{\alpha}{\sigma}\right) k\pi\right]^2} \equiv \sigma M e^{i\beta}.$$

Whence, by integration, we obtain

$$f(x) = M e^{i\beta} \cos\sigma\left(x - x_0 + \frac{\alpha}{\sigma}\right) + C \sin\sigma\left(x - x_0 + \frac{\alpha}{\sigma}\right),$$

where C is an arbitrary constant.

Thus we see that the extremal functions have the form

$$f_0(z) = a e^{i\sigma z} + b e^{-i\sigma z},$$

where a, b are arbitrary complex constants which satisfy the condition

$$|a| + |b| = M,$$

so that we have

$$\sup_{-\infty < x < \infty} |f_0(x)| = M.$$

The proof of the second theorem proceeds in a wholly analogous way.

In the case of a real function $f(x)$ $(-\infty, \infty)$, the theorems just proven take on the following form:

$$\sup_{-\infty<x<\infty} \sqrt{[f'(x)]^2 + \sigma^2 [f(x)]^2} \leqq \sigma \sup_{-\infty<x<\infty} |f(x)|,$$

$$\sup_{-\infty<x<\infty} \sqrt{[f'(x)]^2 + [\tilde{f}'(x)]^2} \leqq \sup_{-\infty<x<\infty} |f(x)|.$$

It follows further from Theorem 2 that $\tilde{f}(z)$ belongs to B_σ whenever $f(z)$ does.

In conclusion, let us present another, somewhat more general, formulation of Bernstein's inequality, which can be obtained immediately from the formulas (I), (II).

THEOREM 3. *Let E be any complete linear normalized space of the function $f(x)$ $(-\infty < x < \infty)$ having the following properties:*

a) *From $f(x) \in E$ there follows $f(z) \in B_\sigma$,*

b) *If $f(x)$ is contained in E and if t is an arbitrary real number, then $f(x+t)$ likewise belongs to E and we have*

$$||f(x+t)|| = ||f(x)||.$$

Then $f'(x)$ and $\tilde{f}'(x)$, as well as $f(x)$, are elements of E and we have

$$(2) \qquad \begin{cases} ||U_1(x)|| \equiv ||\sin\alpha \cdot f'(x) - \sigma\cos\alpha \cdot f(x)|| \leqq \sigma\,||f(x)||, \\ ||U_2(x)|| \equiv ||\sin\alpha \cdot f'(x) + \cos\alpha \cdot \tilde{f}'(x)|| \leqq \sigma\,||f(x)||. \end{cases}$$

Worthy of special attention are the following two spaces which we thus obtain. Let us comprehend all the functions which satisfy condition a) and satisfy the two following inequalities:

$1^0.$ $$\sup_{-\infty<t<\infty} \left\{ \int_t^{2\pi+t} |f(x)|^p\,dx \right\}^{\frac{1}{p}} < \infty \qquad (p \geqq 1),$$

$2^0.$ $$\left\{ \int_{-\infty}^{\infty} |f(x)|^p\,dx \right\}^{\frac{1}{p}} < \infty \qquad (p \geqq 1).$$

As norms in these spaces we take the left-hand sides of 1^0 and 2^0, respectively.

In case 1^0, the functions $e^{i\lambda x}(-\sigma \leq \lambda \leq \sigma)$ are contained in E. Hence in this case the equality sign in the inequalities (2) is taken for those functions of the form

$$f_0(x) = ae^{i\sigma x} + be^{-i\sigma x}$$

which satisfy the equation

$$U_1(x) = U_2(x) = -\sigma f_0\left(x - \frac{\alpha}{\sigma}\right).$$

But also in case 2^0, the inequalities (2) are sharp in the sense that it is impossible to replace the right-hand sides by

$$\sigma \Theta \, || f(x) ||$$

with $\Theta < 1$.

In fact: Suppose we select the functions

$$f_h(x) = e^{i(\sigma-h)x}\left(\frac{\sin\frac{hx}{2}}{\frac{hx}{2}}\right)^2,$$

which obviously belong to the space under consideration for any value of h in $(0, \sigma]$. A simple calculation shows that

$$U_1(x) = -\sigma e^{i\alpha} f_h(x) + h e^{i(\sigma-h)x} \varphi(hx)$$

where $\varphi(x) \in L^p$. Consequently, we have

$$(3) \qquad || U_1(x) || \geqq \sigma || f_h(x) || - h || \varphi(hx) || =$$
$$= \sigma || f_h(x) || - h^{\frac{1}{q}} || \varphi(x) || \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right).$$

On the other hand the equation

$$|| f_h(x) || = \frac{1}{h^{\frac{1}{p}}} \left\| \frac{\sin^2\frac{x}{2}}{\left(\frac{x}{2}\right)^2} \right\|$$

is satisfied. Therefore the ratio of the first term of the right-hand side of the inequality (3) to the second term approaches infinity if h approaches zero. As a result: For every $\varepsilon > 0$ there exists an h such that the inequality

$$|| U_1(x) || > (\sigma - \varepsilon) || f_h(x) ||$$

is valid.

The proposition for the second inequality is established similarly.

To simplify matters, α is put equal to zero. Let us set

$$f_h(x) = \frac{1}{h^3} \int_{\sigma-h}^{\sigma} (\sigma - v)(v - \sigma + h) \cos vx \, dv.$$

From the definition of the conjugate function, we have

$$\tilde{f}_h(x) = -\frac{1}{h^3} \int_{\sigma-h}^{\sigma} (\sigma - v)(v - \sigma + h) \sin vx \, dv.$$

Whence it follows that

$$\tilde{f}'_h(x) = -\sigma f_h(x) + \varphi_h(x),$$

where

$$\varphi_h(x) = \cos \sigma x \frac{6(\cos hx - 1) + 4hx \sin hx - h^2x^2 \cos hx}{h^3x^4} +$$

$$+ \sin \sigma x \frac{6 \sin hx - 2hx - 4hx \cos hx - h^2x^2 \sin hx}{h^3x^4} .$$

Now, since we have

$$\lim_{h \to 0} f_h(x) = -\frac{1}{3!} \cos \sigma x$$

in every bounded interval, uniformly, it follows that

$$\lim_{h \to 0} || f_h(x) || = \infty .$$

On the other hand, however, the quantity

$$|| \varphi_h(x) ||$$

remains bounded as h approaches zero. Therefore, when $h > 0$ is small enough, we get

$$|| \tilde{f}_h'(x) || > (\sigma - \varepsilon) || f_h(x) ||,$$

no matter how small $\varepsilon > 0$ may be chosen. This proves the theorem.

75. B. M. Levitan's Polynomials [60]. Suppose given a function $f(x)$ $(-\infty < x < \infty)$ which fulfills the following conditions:

a) $f(x) \in W^2$, that is

$$\frac{f(x)}{1 + |x|} \in L^2(-\infty, \infty),$$

b) $f(x)$ is the same as an integral transcendental function $f(z)$ of the exponential type when the latter is restricted to the real axis.

Let us introduce into our investigation the integral

$$E_h(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \left(\frac{2 \sin \frac{hu}{2}}{hu} \right)^2 f(u)\, du \qquad (-\infty < x < \infty),$$

which is dependent on a positive parameter h. Let us assume that $f(x)$ belongs to E_τ $(\sigma \geqq \tau)$ and let us set $h = \frac{\sigma}{n}$, where n is a positive integer. With M. G. Krein [59] we will call the trigonometric sum

$$S_n(f; x) = \sum_{k=-n}^{\infty} hE_h(kh)\, e^{ikhx}$$

a *Levitan Polynomial* of the function $f(x)$.

When the function is real-valued, so are its Levitan polynomials. Furthermore, the Levitan polynomials possess the following properties:

1′ From $f(x) \equiv 1$ follows $S_n(f; x) \equiv 1$,

2′ From $f(x) \geq 0$ $(-\infty < x < \infty)$ follows $S_n(f; x) \geq 0$ $(-\infty < x < \infty)$,

3′ From $|f(x)| \leq M (-\infty < x < \infty)$ follows $|S_n(f; x)| \leq M (-\infty < x < \infty)$.

Property 1′ results from the fact that the equation

$$(1) \qquad E_h(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \left(\frac{2 \sin \frac{hu}{2}}{hu} \right)^2 du = \begin{cases} \frac{1}{h} \left(1 - \frac{|x|}{h} \right) & (-h \leqq x \leqq h) \\ 0 & (|x| > h) \end{cases}$$

holds when $f(x) \equiv 1$.

Property 3′ follows from 2′ and 1′.

In fact: The validity of the inequality

$$|\varphi(x)| \leqq M \qquad (-\infty < x < \infty)$$

is equivalent to the statement that, for an arbitrary real α, we have

$$-M \leqq \Re \{e^{i\alpha} \varphi(x)\} \leqq M \qquad (-\infty < x < \infty) .$$

Hence, if

$$|f(x)| \leqq M ,$$

that is, if

$$-M \leqq \Re \{e^{i\alpha} f(x)\} \leqq M ,$$

it follows from 1′ and 2′ that

$$-M \leqq \Re \{e^{i\alpha} S_n(f; x)\} \leqq M$$

and, accordingly, also that

$$|S_n(f; x)| \leqq M .$$

Property 2′ will be established later.

Now, B. M. Levitan has proven the following theorem: *If $f(z)$ is contained in B_σ, then we have*

$$\lim_{n \to \infty} S_n(f; x) = f(x) \qquad (-\infty < x < \infty)$$

uniformly in every bounded interval.

M. G. Krein has given a new proof of Levitan's theorem. This enables us to estimate the rapidity with which the sequence of the polynomials $S_n(f; x)$ converges to the function $f(x)$. With the aid of Krein's method, we are going to prove the following, somewhat more general theorem.

THEOREM: *If the function* $f(x)$ $(-\infty < x < \infty)$ *admits a representation*

$$f(x) = f(0) + x\int_{-\sigma}^{\sigma} e^{iux}\,\psi(u)\,du \tag{2}$$

where $\psi(u) \in L^2(-\sigma, \sigma)$, *then we have*

$$|S_n(f;x) - f(x)| \leqq \frac{\sigma x^2}{n}\int_{-\sigma}^{\sigma} |\psi(u)|\,du. \tag{3}$$

Since every function in B_σ has a representation of the above form, Levitan's theorem is a special case of the theorem just formulated.

If $f(z)$ is contained in B_σ, then $\int_{-\sigma}^{\sigma} \psi(u)\,|\,du\,|$ can readily be estimated with the aid of

$$M = \sup_{-\infty<x<\infty} |f(x)|\,.$$

In fact, we have

$$2\pi\int_{-\sigma}^{\sigma} |\psi(u)|^2\,du = \int_{-\infty}^{\infty} \frac{|f(x)-f(0)|^2}{x^2}\,dx =$$

$$= \int_{-\frac{2}{\sigma}}^{\frac{2}{\sigma}} \frac{|f(x)-f(0)|^2}{x^2}\,dx + \int_{x>\frac{2}{\sigma}} \frac{|f(x)-f(0)|^2}{x^2}\,dx \leqq$$

$$\leqq M^2\sigma^2 \cdot \frac{4}{\sigma} + M^2 \cdot 4\sigma = 8M^2\sigma.$$

After applying Schwarz's inequality, we find

$$\int_{-\sigma}^{\sigma} |\psi(u)|\,du \leqq \frac{1}{\sqrt{2\pi}}\sqrt{2\sigma}\sqrt{8M^2\sigma} = \frac{4M\sigma}{\sqrt{2\pi}}.$$

With the aid of the theorem on which the proof is based, we now have

$$|S_n(f;x) - f(x)| \leqq \frac{4}{\sqrt{2\pi}} M\sigma^2\frac{x^2}{n} \quad (-\infty < x < \infty), \tag{4}$$

$$(5) \qquad |S_n(\tilde{f}; x) - \tilde{f}(x)| \leqq \frac{4}{\sqrt{2\pi}} M\sigma^2 \frac{x^2}{n} \quad (-\infty < x < +\infty),$$

provided $f(z)$ lies in B_σ and $M = \sup_{-\infty<x<\infty} |f(x)|$.

We now proceed to the proof of inequality (3); let us observe that, by virtue of property 1', the equation

$$S_n(f; x) = f(0) + S_n(f_1; x)$$

is true if we set

$$f_1(x) = f(x) - f(0) .$$

On the other hand, the function $f_1(x)$ satisfies the equation

$$E_h(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{-iux} \frac{1 - \cos hu}{h^2 u} \left\{ \int_{-\sigma}^{\sigma} e^{itu} \psi(t)\, dt \right\} du .$$

By virtue of the theorem on the Faltung of functions (cf. Section 58), this equation can be put in the following form

$$(6) \qquad E_h(x) = \frac{i}{h^2} \int_{x}^{x+h} \psi(t)\, dt - \frac{i}{h^2} \int_{x-h}^{x} \psi(t)\, dt ,$$

if we assume that $\psi(t)$, though defined along the entire real axis, is nevertheless equal to zero when $t > \sigma$ and $t < -\sigma$.

The foregoing shows, among other things, that

$$(7) \qquad E_h(x) = 0 \qquad (|x| \geqq \sigma + h) .$$

Because of (1), this equation is true also for $f(x)$.

We can now give an expression for $S_n(f; x)$ which we shall need subsequently. This has the form

$$(8) \qquad S_n(f; x) = f(0) + \sum_{k=-n}^{n} \frac{i}{h} e^{ikhx} \left\{ \int_{kh}^{(k+1)h} \psi(t)\, dt - \int_{(k-1)h}^{kh} \psi(t)\, dt \right\} =$$

$$= f(0) + \sum_{k=-n}^{n-1} \frac{e^{i(k+1)hx} - e^{ikhx}}{hi} \int_{kh}^{(k+1)h} \psi(t)\, dt .$$

Since, however,

$$f(x) = f(0) + x \sum_{k=-n}^{n-1} \int_{kh}^{(k+1)h} e^{itx} \psi(t)\, dt$$

it follows that

$$S_n(f;x)-f(x)=\sum_{k=-n}^{n-1}\int_{kh}^{(k+1)h}\psi(t)\left\{\frac{e^{i(k+1)hx}-e^{ikhx}}{hi}-xe^{itx}\right\}dt.$$

Hence we have

$$S_n(f;x)-f(x)=x\sum_{k=-n}^{n-1}\int_{kh}^{(k+1)h}\psi(t)\,(e^{ikhx}-e^{itx})\,dt$$

with $kh<t_k<(k+1)h$. Whence we now obtain the estimate:

$$|S_n(f;x)-f(x)|\leqq x^2h\sum_{k=-n}^{n-1}\int_{kh}^{(k+1)h}|\psi(t)|\,dt=\frac{\sigma x^2}{n}\int_{-\sigma}^{\sigma}|\psi(u)|\,du.$$

This proves the theorem.

It will be observed further that there exist functions $f(z)$ in B_σ such that the equations $\tilde{f}(0)=0$, $\tilde{S}_n(f;0)=0$,

$$\tilde{S}_n(f;x)=S_n(\tilde{f};x)$$

are valid. Therefore, by applying our usual method of normalization, we obtain from (5) the following inequality:

$$|\tilde{S}_n(f;x)-\tilde{f}(x)|\leqq\frac{4}{\sqrt{2\pi}}M\sigma^2\frac{x^2}{n}.$$

We will now prove property 2′. To that end, let us bring $S_n(f;x)$ into the form

$$S_n(f;x)=\lim_{N\to\infty}\sum_{k=-N}^{N}h\left(1-\frac{|k|}{N}\right)E_h(kh)\,e^{ikhx}=$$

$$=\lim_{N\to\infty}\sum_{k=-N}^{N}h\left(1-\frac{|k|}{N}\right)E_h(kh)\,e^{ikhx},$$

which is permissible by virtue of (7). If we now keep in mind also that

$$\sum_{k=-N}^{N}h\left(1-\frac{|k|}{N}\right)E_h(kh)\,e^{ikhx}=\frac{2}{\pi}\int_{-\infty}^{\infty}f(u)\,\frac{\sin^2\frac{hu}{2}}{hu^2}\sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)e^{ikh(x-u)}\,du=$$

$$=\frac{2}{\pi N}\int_{-\infty}^{\infty}f(u)\,\frac{\sin^2\frac{hu}{2}}{hu^2}\left|\sum_{k=0}^{N-1}e^{ikh(x-u)}\right|^2du,$$

then the proof is complete.

In conclusion, let us take up also the case in which the function $f(z)$ in B_σ is of the form

$$f(z) = \int_{-\sigma}^{\sigma} e^{iuz} d\omega(u) .$$

Here $\omega(u)$ is a function of bounded variation. We are now going to prove that in this case the equation

$$\lim_{n\to\infty} \sum_{k=-n}^{n} hE_h(kh)\, \lambda(kh) = \int_{-\sigma}^{\sigma} \lambda(u)\, d\omega(u)$$

holds for every continuous function $\lambda(u)$ $(-\sigma \leq u \leq \sigma)$.

In order to do so, let us write $f(x)$ in the form

$$f(x) = \alpha e^{i\sigma x} + \beta e^{-i\sigma x} + \frac{x}{i} \int_{-\sigma}^{\sigma} e^{iux}\, \omega(u)\, du .$$

Our proposition can be verified immediately for the function

$$f(x) = e^{\pm i\sigma x} .$$

In fact: As we have seen above (cf. 1), we have in this case

$$E_h(x) = \begin{cases} \frac{1}{h}\left\{1 - \frac{|\pm\sigma - x|}{h}\right\} & (-h \leqq \pm\sigma - x \leqq h), \\ 0 & (|\pm\sigma - x| > h), \end{cases}$$

but that implies

$$\sum_{k=-n}^{n} hE_h(kh)\, \lambda(kh) = \lambda(\pm\sigma) .$$

Therefore $f(x)$ can be put in the special form

$$f(x) = \frac{x}{i} \int_{-\sigma}^{\sigma} e^{iux}\, \omega(u) du .$$

But this expression for $f(x)$ coincides with that previously given if we set

$$f(0) = 0, \qquad \omega(u) = i\psi(u) .$$

It follows then from (6) that

$$E_h(x) = \frac{1}{h^2} \int_{x}^{x+h} \omega(t)\, dt - \frac{1}{h^2} \int_{x-h}^{h} \omega(t)\, dt =$$

$$= \frac{1}{h}\left\{\omega(x+h) - \omega(x)\right\} - \frac{1}{h^2} \int_{x}^{x+h} (t-x)\, d\omega(t) - \int_{x-h}^{x} (t - x + h)\, d\omega(t)\Big\}.$$

Whence it follows that

$$\sum_{k=-n}^{n} hE_h(kh)\,\lambda(kh) = \sum_{k=-n}^{n} \lambda(kh)\,\{\omega([k+1]h) - \omega(kh)\} -$$

$$- \sum_{k=-n}^{n} \lambda(kh)\,\frac{1}{h}\left\{\int_{kh}^{(k+1)h} (t - kh)\,d\omega(t) - \int_{(k-1)h}^{kh} [t - (k-1)h]\,d\omega(t)\right\}.$$

As $n \to \infty$ the first term of the right-hand side converges to

$$\int_{-\sigma}^{\sigma} \lambda(u)\,d\omega(u).$$

Hence it remains merely to show that the second term approaches zero. Obviously this term is equal to

$$I_n = \sum_{k=-(n-1)}^{n} \{\lambda(kh) - \lambda([k-1]h)\}\,\frac{1}{h}\int_{(k-1)h}^{kh} [t - (k-1)h]\,d\omega(t).$$

Now let

$$\varepsilon_n = \max_k |\lambda(kh) - \lambda([k-1]h)|.$$

Then we have

$$|I_n| \leqq \varepsilon_n \sum_{k=-(n-1)}^{n} \int_{(k-1)h}^{kh} |d\omega(t)| = \varepsilon_n \operatorname{Var} \omega(t).$$

But since $\lim_{n\to\infty} \varepsilon_n = 0$, there follows $I_n \to 0$, as we were required to prove.

76. The Theorem of Fejér and Riesz. A Generalization of This Theorem.

Let

$$f(z) = \sum_{k=-n}^{n} A_k e^{ikz} \qquad (z = x + iy)$$

be a trigonometric sum which assumes only real non-negative values on the real axis. Then there exists a polynomial

$$g(z) = \sum_{k=0}^{n} B_k e^{ikz}$$

such that

$$f(x) = |g(x)|^2 \qquad (-\infty < x < \infty).$$

The polynomial $g(x)$ can be so constructed that all of its roots are situated in the half-plane $y \geqq 0$.

This theorem is due to Fejér and Riesz [61]. For the proof, let us set

$$e^{iz} = \zeta,$$

hence

$$f(z) = \sum_{k=-n}^{n} A_k \zeta^k = F(\zeta) .$$

Since the function $F(\zeta)$ is real-valued along the circle $|\zeta| = 1$ it follows that $F(\zeta)$ satisfies the equation

$$F(\zeta) = F\left(\frac{1}{\bar{\zeta}}\right).$$

Hence all of its roots are situated symmetrically with respect to the circle $|\zeta| = 1$.

Now, let

$$\zeta_1, \zeta_2, \ldots, \zeta_r$$

be all the roots of $F(\zeta)$ which lie in the interior of the circle $|\zeta| < 1$. Then $F(\zeta)$ will have the roots

$$\frac{1}{\bar{\zeta}_1}, \frac{1}{\bar{\zeta}_2}, \ldots, \frac{1}{\bar{\zeta}_r}$$

in the region $|\zeta| > 1$. Suppose $A_n \neq 0$, so that $F(\zeta)$ has exactly $2n$ roots. Then $2n - 2r$ roots

$$\Theta_1, \Theta_2, \ldots, \Theta_{2n-2r}$$

of $F(\zeta)$ will lie along the circle $|\zeta| = 1$. Accordingly, we have

$$F(\zeta) = \frac{A_n}{\zeta^n} \prod_{k=1}^{r} (\zeta - \zeta_k)\left(\zeta - \frac{1}{\bar{\zeta}_k}\right) \prod_{m=1}^{2n-2r} (\zeta - \Theta_m) =$$

$$= \left\{ \prod_{k=1}^{r} (\zeta - \zeta_k)\left(\frac{1}{\zeta} - \bar{\zeta}_k\right) \right\} \left\{ \frac{C}{\zeta^{n-r}} \prod_{m=1}^{2n-2r} (\zeta - \Theta_m) \right\}.$$

The first factor is real and non-negative along the circle $|\zeta| = 1$. Therefore the second factor must also possess these properties. Consequently the quantities Θ_m must be equal pairwise. Hence the second factor has the form

$$A \prod_{m=1}^{s} (\zeta - \Theta_m)\left(\frac{1}{\zeta} - \bar{\Theta}_m\right).$$

Obviously, here A must be positive. If we now further set

$$g(z) = \sqrt{A} \prod_{k=1}^{r} (\zeta - \zeta_k) \prod_{m=1}^{s} (\zeta - \Theta_m),$$

our theorem is proved.

Furthermore, the following general statement is true: *If* $f(z)$ *is contained in* B_σ, *and if* $f(x) \geqq 0$ $(-\infty < x < \infty)$, *there exists in* $B_{\frac{\sigma}{2}}$ a *function* $g(z)$ *such that*

$$f(x) = |g(x)|^2 \qquad (-\infty < x < \infty)$$

where $g(z)$ *can be so chosen that all of its zeros lie in the half-plane* $y \geqq 0$.

With M. G. Krein, we prove this theorem on the basis of the Fejér-Riesz theorem with the aid of a passage to a limit. To that end, let us construct the Levitan polynomials, which, by virtue of the hypothesis in the theorem to be proved, are non-negative:

$$S_n(f; x) \geqq 0 \qquad (-\infty < x < \infty).$$

According to the theorem of Fejér and Riesz, we then have

$$S_n(f; x) = \left| \sum_{k=0}^{n} B_k^{(n)} e^{ikhx} \right|^2 = |g_n(x)|^2$$

with

$$g_n(z) = e^{-\frac{i\sigma z}{2}} \sum_{k=0}^{n} B_k^{(n)} e^{\frac{ik\sigma z}{n}}.$$

Let us now assume that all the roots of $g_n(z)$ are to be found in the half-plane $y \geqq 0$. Since $g_n(z)$ belongs to $B_{\frac{\sigma}{2}}$ and since the inequality

$$|g_n(x)| \leqq \sqrt{M} \qquad (-\infty < x < \infty),$$

where

$$M = \sup_{-\infty < x < \infty} |f(x)|$$

is valid, we have a uniformly bounded sequence of functions along the real axis, if n runs through the values 1, 2, 3, ... Now we need only recall the principle of compactness (cf. Section 73) and apply Levitan's theorem, by reason of which

$$S_n(f; x) \rightrightarrows f(x).$$

The assumption that all the roots of $g(z)$ are to be found in the half-plane $y \geqq 0$ follows from the fact that the functions $g_n(z)$ have no roots in the half-plane $y < 0$.

77. A Criterion for the Representation of Continuous Functions as Fourier-Stieltjes Integrals. The extremum problem broached in Section 73 pertained to the trigonometric sums transformed by means of a given sequence of multipliers. When this problem is formulated in its generality, it involves the question of representing a given continuous function $\lambda(t)$ $(-\sigma \leq t \leq \sigma)$ as a Fourier-Stieltjes integral

$$\lambda(t) = \int_{-\infty}^{\infty} e^{iut}\, d\omega(u) \ .$$

In this connection, $\omega(t)$ is assumed to be a function of bounded variation.

In the recent work by M. G. Krein this problem is studied exhaustively. He proved [59] the equivalence of the following two statements:

a) Every continuous function $\lambda(t)$ $(-\sigma \leq t \leq \sigma)$ has a representation of the form

$$\lambda(t) = \int_{-\infty}^{\infty} e^{iut}\, d\omega(u) \qquad (-\sigma \leqq t \leqq \sigma) \tag{1}$$

where

$$\operatorname{Var} \omega(u) \leqq M .$$

b) The inequality

$$\left| \sum_{k=1}^{n} c_k \lambda(t_k) \right| \leqq M \sup_{-\infty < x < \infty} \left| \sum_{k=1}^{n} c_k e^{it_k x} \right| \tag{2}$$

is valid, where n is a positive integer, t_k is real and contained in $[-\sigma, \sigma]$ and c_k $(k = 1, 2, \ldots, n)$ may be complex.

Proof: It is easy to verify that b) follows from a). For, in the case of a function of the form (1) we have

$$\sum_{k=1}^{n} c_k \lambda(t_k) = \int_{-\infty}^{\infty} \sum_{k=1}^{n} c_k e^{it_k u}\, d\omega(u) .$$

Whence it follows that

$$\left| \sum_{k=1}^{n} c_k \lambda(t_k) \right| \leqq \sup_{-\infty < u < \infty} \left| \sum_{k=1}^{n} c_k e^{it_k u} \right| \cdot \operatorname{Var} \omega(u) .$$

To prove the converse, we introduce the space $\mathfrak{M}_\sigma \subset C_\infty$ of all functions of the form

$$f(x) = \int_{-\sigma}^{\sigma} e^{itx}\, \varphi(t)\, dt \tag{3}$$

with $\varphi(t) \in L(-\sigma, \sigma)$. The integral

$$f_\lambda = \int_{-\sigma}^{\sigma} \lambda(t)\, \varphi(t)\, dt \tag{4}$$

is obviously a homogeneous, additive functional on $\mathfrak{M}_\sigma$. With the aid of (2), we will now show that this functional is bounded, namely, that

$$|f_\lambda| \leqq M \max_{-\infty \leqq x \leqq \infty} |f(x)| = M\, ||f|| . \tag{5}$$

For that purpose, let us construct the Levitan polynomial $S_n(f; x)$. By virtue of (2) and the theorem of Levitan, we then have

$$\left| \sum_{k=-n}^{n} hE_h(kh)\, \lambda(kh) \right| \leqq M \sup_{-\infty<x<\infty} |S_n(f; x)| \leqq M\, ||f||.$$

Whence we obtain the inequality (5), inasmuch as (cf. Section 75)

$$\lim_{n\to\infty} \sum_{k=-n}^{n} hE_h(kh)\, \lambda(kh) = f_\lambda .$$

Hence f_λ is a linear functional on $\mathfrak{M}_\sigma$ having a finite norm which does not exceed M. Therefore the functional can be extended over the whole of space C_∞ without increasing the norm. Now, the following is valid: Every linear functional in C_∞ is of the form

$$\Phi(f) = uf(\infty) + \int_{-\infty}^{\infty} f(x)\, d\omega(x)$$

with

$$|\mu| + \operatorname*{Var}_{-\infty<x<\infty} \omega(x) = ||\Phi||\ ^{4}$$

Wherefore there exists a function $\omega(x)$ such that

$$f_\lambda = \int_{-\infty}^{\infty} f(x)\, d\omega(x) \qquad (f \in \mathfrak{M}_\sigma) \tag{6}$$

and

$$\operatorname*{Var}_{-\infty<x<\infty} \omega(x) \leqq M .$$

Here it must be remembered that every function $f(x)$ in $\mathfrak{M}_\sigma$ vanishes at infinity.

If we now combine the formulas (3), (4) and (6), we obtain the equation

$$\int_{-\sigma}^{\sigma} \lambda(t)\, \varphi(t)\, dt = \int_{-\infty}^{\infty} \left\{ \int_{-\sigma}^{\sigma} e^{itu}\, \varphi(u)\, du \right\} d\omega(t) = \int_{-\sigma}^{\sigma} \varphi(u) \left\{ \int_{-\infty}^{\infty} e^{itu}\, d\omega(t) \right\} du .$$

As $\varphi(t)$ may be chosen arbitrarily in $L(-\sigma, \sigma)$, the equation

$$\lambda(t) = \int_{-\infty}^{\infty} e^{itu}\, d\omega(u) \tag{7}$$

holds everywhere in $[-\sigma, \sigma]$. But since both sides of this equation are continuous (the left-hand side is continuous in $[-\sigma, \sigma]$, by hypothesis; the right-hand side by reason of its structure, even along the entire real axis), the equation (7) is valid, not only almost everywhere in $[-\sigma, \sigma]$, but throughout the interval. Thus the theorem is proved.

[4] A proof of this will be given at the end of this section.

This theorem holds also in the case where, instead of a closed interval $[-\sigma, \sigma]$, a bounded open interval or the infinite interval $(-\infty, \infty)$ is considered. For the case of the infinite interval $(-\infty, \infty)$ the theorem was first established by Bochner [62].

Let us now turn to the special case in which

$$\lambda(t) = \varepsilon \int_{-\infty}^{\infty} e^{-iu\sigma} e^{itu} d\omega_0(u) \qquad (-\sigma \leqq t \leqq \sigma;\ |\varepsilon| = 1) \tag{8}$$

where $\omega_0(u)$ is a non-decreasing function of bounded variation. (It is easy to see that $\operatorname{Var} \omega_0(u) = |\lambda(\sigma)|$). In this case $|\lambda(\sigma)|$ is the least value of the constant M in relation (2). As M. G. Krein has shown, this is a necessary condition that $\lambda(t)$ can be represented in the form (8).

Let us now assume that $\lambda(t)$ has the form (8). The question now arises: For which of the functions $f(z)$ in B_σ, representable in the form

$$f(z) = \int_{-\sigma}^{\sigma} e^{izt} d\Omega(t) \tag{9}$$

is the equality sign admitted in the relation

$$\sup_{-\infty < x < \infty} |f_\lambda(x)| \leqq |\lambda(\sigma)| \sup_{-\infty < x < \infty} |f(x)| \tag{10}$$

where

$$f_\lambda(z) = \int_{-\sigma}^{\sigma} e^{izt} \lambda(t)\, d\Omega(t)\ ?$$

Evidently the expressions $Ce^{i\sigma z}$ are always "extremal functions." However, these are not always the only ones. In fact: By reason of the relations (8), (9) we have

$$f_\lambda(x) = \varepsilon \int_{-\infty}^{\infty} e^{-iu\sigma} f(x+u)\, d\omega_0(u).$$

Hence the function $f(z)$ is an extremal if and only if, for any value of x (which we will call x_0), the product

$$e^{-iu\sigma} f(x_0 + u)$$

assumes one and the same value at every point at which the function $\omega_0(u)$ increases and is subject to the condition

$$\sup_{-\infty < x < \infty} |f(x)| = M.$$

This is certainly true in the trivial case where $f(z)$ is equal to $Ce^{i\sigma z}$. In order that there shall exist also a non-trivial case, it is necessary (by reason of a property of analytic functions)

that $\omega_0(u)$ be a pure step function. In addition, the jumps must occur at the zeros of the function

$$F(u) = M^2 - f(x_0 + u)\overline{f(x_0 + u)}.$$

Obviously the function $F(z)$ lies in $B_{2\sigma}$ and from $F(u) \geq 0$ $(-\infty < u < \infty)$ it follows that

$$F(u) = |G(u)|^2$$

with $G(z) \in B_\sigma$.

In order that the equality sign be admissible in the relation (10) for the non-trivial case, the function $\omega_0(u)$ in the representation (8) must therefore be a non-decreasing step function whose points of discontinuity are zeros of a certain function in B_σ. A good illustration of the foregoing is furnished by our proof of Bernstein's inequality (cf. Section 74). In fact, we obtained therein representations of the form (8) for the functions $\omega_0(u)$. The functions $V_1(t)$, $V_2(t)$, respectively, were in both instances step functions of the kind mentioned above.

In conclusion, let us take up the proof, which was deferred earlier in this section, of the assertion concerning the form of the linear functionals in C_∞.

It suffices to show that every linear functional $\Phi(f)$ in the space C_π of all continuous functions of period 2π is of the form

$$\Phi(f) = \mu f(\pi) + \int_{-\pi+0}^{\pi-0} f(t)\, d\omega(t)$$

with

$$||\Phi|| = |\mu| + \operatorname*{Var}_{-\pi<t<\pi} \omega(t).$$

If we extend this functional to the space $C(-\pi, \pi)$ of all the continuous functions in the interval $[-\pi, \pi]$, without enlarging the norm, then, by reason of the theorem on the form of linear functionals in the space C (cf. Section 29, where the space $C(0, 1)$ was considered), the extended linear functional has a representation

$$\Phi(g) = \int_{-\pi}^{\pi} g(t)\, d\omega(t) \qquad [g \in C(-\pi, \pi)]$$

with

$$\operatorname*{Var}_{-\pi\leq t\leq\pi} \omega(t) = ||\Phi|| = ||\Phi||_{\mathcal{C}_\pi}.$$

This can be put into the form

$$\Phi(g) = g(\pi)[\omega(\pi) - \omega(\pi - 0)] + g(-\pi)[\omega(-\pi + 0) - \omega(-\pi)] + \int_{-\pi+0}^{\pi-0} g(t)\, d\omega(t)$$

with

$$|\omega(\pi) - \omega(\pi - 0)| + |\omega(-\pi + 0) - \omega(-\pi)| + \operatorname*{Var}_{-\pi<t<\pi} \omega(t) = ||\Phi||.$$

Now, if $f(x)$ is contained in C_π, then $f(\pi) = f(-\pi)$ and consequently

$$\Phi(f) = \mu f(\pi) + \int_{-\pi+0}^{\pi-0} f(t)\, d\omega(t)$$

with

$$\mu = \omega(\pi) - \omega(\pi - 0) + \omega(-\pi + 0) - \omega(-\pi).$$

But since

$$||\Phi||_{C_\pi} \leqq |\mu| + \operatorname*{Var}_{-\pi<t<\pi} \omega(t) \leqq |\omega(\pi) - \omega(\pi - 0)| + |\omega(-\pi + 0) - \omega(-\pi)| + \operatorname*{Var}_{-\pi<t<\pi} \omega(t) = ||\Phi|| = ||\Phi||_{C_\pi},$$

we have

$$||\Phi||_{C_\pi} = |\mu| + \operatorname*{Var}_{-\pi<t<\pi} \omega(t).$$

CHAPTER V

QUESTIONS REGARDING THE BEST HARMONIC APPROXIMATION OF FUNCTIONS

78. Preliminary Remarks. We already know that, to every continuous periodic function $f(x)$ and every positive integer n, there corresponds a completely defined trigonometric sum — denoted here by $\mathcal{T}_{n-1}(x;f)$ — which, among all the trigonometric sums of the order $< n$, deviates the least from $f(x)$. The measure of the deviation, the function

$$\varrho_n = \max_{0 \leq x \leq 2\pi} |f(x) - \mathcal{T}_{n-1}(x;f)|$$

is then a functional of $f(x)$ dependent on n. By reason of Weierstrass' theorem, we then have

$$\lim_{n \to \infty} \varrho_n = 0.$$

However, this theorem tells us nothing at all about the rapidity with which ϱ_n approaches zero as $n \to \infty$. Now, it is entirely natural to expect the rapidity to depend on the differential properties of $f(x)$; thus, for example, on the existence of a bounded derivative of one order or another, on whether or not $f(x)$ is analytic, etc.

Lebesgue [63] and de la Vallée-Poussin [64] proposed the first problem relating thereto. They were concerned with the approximation through polynomials and not through trigonometric sums. De la Vallée-Poussin worded it as follows: "It would be interesting to know whether or not it is possible to represent the ordinate of a broken line approximately by means of polynomials of the n-th degree with a measure of accuracy better than the order $\frac{1}{n}$." This question has been fully answered by S. N. Bernstein [65].

Now, thanks to the efforts of D. Jackson [66] and S. N. Bernstein [65], [67], [68], the connection between the differential properties of $f(x)$ and the rate of decrease in ϱ_n has been investigated quite thoroughly.

Analogous questions arise also when dealing with approximation in other spaces, for instance in $L^p(0,2\pi)$.

We need not restrict ourselves necessarily to periodic functions; we can include also more comprehensive manifolds of functions defined along the entire real axis, e.g. the totality of all bounded functions, of all functions in $L^p(-\infty, \infty)$, etc. If we so extend the domain of functions sought to be approximated, then the class of all the trigonometric sums of the order $< T$ is no longer suitable to

serve as a set of approximating functions. To suit the circumstances, we replace the trigonometric sums by the more comprehensive class [1] G_{T-0} [69]. G_{T-0} consists of those functions which are identical with any integral transcendental function of the exponential type with an index $< T$ and restricted to the real axis.

Since, in the class G_{T-0}, the functions

$$\int_{-T}^{T} e^{iux}\varphi(u)\,du, \quad C + x\int_{-T}^{T} e^{iux}\psi(u)\,du,$$

that is trigonometric integrals, are of prime importance, we have the right to speak of "harmonic approximation" (in the ordinary or in the generalized sense), the theory of which may be regarded as a chapter in general harmonic analysis. We begin this chapter with certain functional-theoretic considerations and then follow with problems of approximation.

79. The Modulus of Continuity. Let $f(x)$ be a function uniformly continuous in a finite or an infinite interval I. Then the function

$$\omega(\delta) = \omega(\delta; f) = \sup_{|x''-x'|\leq\delta} |f(x'') - f(x')| \quad (x', x'' \in I)$$

is called its *modulus of continuity*. Obviously $\omega(\delta)$ is a non-decreasing continuous function of δ which approaches zero with δ.

The validity of the inequality (m a positive integer)

$$\omega(m\delta) \leqq m\omega(\delta)$$

follows from the equation

$$f(x+mh) - f(x) = \sum_{k=0}^{m-1} \{f(x+[k+1]h) - f(x+kh)\}.$$

Furthermore,

$$\omega(\lambda\delta) \leqq (\lambda+1)\,\omega(\delta)$$

is satisfied for any positive λ.

If the above integer m satisfies the inequality

$$m-1 < \lambda \leqq m,$$

we have

$$\omega(\lambda\delta) \leqq \omega(m\delta) \leqq m\omega(\delta) \leqq (\lambda+1)\,\omega(\delta).$$

---

[1] To distinguish between the classes of integral transcendental functions of the exponential type, we will employ the notation T or $T-0$, depending on whether we are speaking of such functions with an exponent $\leq T$ or of such with an exponent $< T$.

In addition, let us prove that $f(x)$ is a constant if the relation

$$\lim_{\delta \to 0} \frac{\omega(\delta)}{\delta} = 0$$

holds; now, by reason of what we have proved above, we have

$$\omega(\delta) = \omega\left(m \frac{\delta}{m}\right) \leqq m\omega\left(\frac{\delta}{m}\right) = \delta \frac{\omega(\delta_1)}{\delta_1} \qquad \left(\delta_1 = \frac{\delta}{m}\right).$$

Hence if

$$\lim_{\delta_1 \to 0} \frac{\omega(\delta_1)}{\delta_1} = 0,$$

then $\omega(\delta)$ is equal to zero whenever δ is positive. But that is impossible unless $f(x)$ is a constant.

We say that a function $f(x)$ fulfills the Lipschitz condition of order α and we write in such case $f(x) \in \mathrm{Lip}\,\alpha$, if $\omega(\delta)$ satisfies the inequality $\omega(\delta) \leq M\delta^{\alpha}$ $(0 < \alpha \leq 1)$, where M is a certain constant. (α must not be greater than unity unless $f(x)$ is a constant.)

If, however,

$$\lim_{\delta \to 0} \omega(\delta) \ln \frac{1}{\delta} = 0,$$

we say that $f(x)$ fulfills the Dini-Lipschitz condition.

It is oftentimes more advantageous to employ, instead of $\omega(\delta)$, the following function:

$$\omega^*(\delta) = \omega^*(\delta; f) = \sup_{|x''-x'| \leq \delta} \left| f(x') + f(x'') - 2f\left(\frac{x'+x''}{2}\right) \right| \quad (x', x'' \in I).$$

With Zygmund [70], we say that $f(x)$ satisfies the condition Λ^* [or λ^*] and we write $f(x) \in \Lambda^*$ [or $f(x) \in \lambda^*$] if $\omega^*(\delta) \leq M(\delta)$ [or $\lim_{\delta \to 0} \frac{\omega^*(\delta)}{\delta} = 0$].

If $f(x)$ is contained in Lip 1, then it belongs also to Λ^*. The converse, however, is not true [see further under Section 81].

80. The Generalization to the Space L^p $(p \geq 1)$. If $f(x)$ lies in the space $L^p(-\infty, \infty)$, we set

$$\omega_p(\delta) = \omega_p(\delta; f) = \sup_{0 < h \leq \delta} \left\{ \int_{-\infty}^{\infty} |f(x+h) - f(x)|^p \, dx \right\}^{\frac{1}{p}}.$$

Then $\omega_p(\delta)$ is also a non-decreasing function of δ.

Let us now derive the equation

$$\lim_{\delta \to 0} \omega_p(\delta) = 0. \tag{1}$$

Let ε be an arbitrary preassigned positive number, and let us determine a corresponding $N > 1$ such that

$$\int_{|x|>N-1} |f(x)|^p \, dx \leqq \left(\frac{\varepsilon}{4}\right)^p$$

We may assume $\delta < 1$; then we have

$$\omega_p(\delta) \leqq \sup_{0<h\leqq\delta} \left\{ \int_{-N}^{N} |f(x+h) - f(x)|^p \, dx \right\}^{\frac{1}{p}} + \frac{\varepsilon}{2}.$$

Let us now construct (cf. Section 24) a polynomial $\Phi(x)$ which satisfies the inequality

$$\left\{ \int_{-N-1}^{N+1} |f(x) - \Phi(x)|^p \, dx \right\}^{\frac{1}{p}} \leqq \frac{\varepsilon}{4} \quad .$$

Then we have:

$$\left\{ \int_{-N}^{N} |f(x+h) - f(x)|^p \, dx \right\}^{\frac{1}{p}} = \left\{ \int_{-N}^{N} |f(x+h) - \Phi(x+h) + \Phi(x+h) -$$

$$- \Phi(x) + \Phi(x) - f(x)|^p \, dx \right\}^{\frac{1}{p}} \leqq \frac{\varepsilon}{4} + \left\{ \int_{-N}^{N} |\Phi(x+h) - \Phi(x)|^p \, dx \right\}^{\frac{1}{p}} + \frac{\varepsilon}{4} \leqq$$

$$\leqq \frac{\varepsilon}{2} + (2N)^{\frac{1}{p}} \cdot \max_{\substack{-N\leqq x\leqq N \\ 0<h\leqq\delta}} |\Phi(x+h) - \Phi(x)|.$$

But since the right-hand side converges to $\frac{\varepsilon}{2}$ for $\delta \to 0$, there follows

$$\overline{\lim_{\delta \to 0}} \omega_p(\delta) \leqq \varepsilon,$$

and, as $\varepsilon > 0$ can of course be chosen entirely arbitrarily, our proposition is proved.

From (1) and the Minkowski inequality, it follows that $\omega_p(\delta)$ is a continuous function of δ. We call $\omega_p(\delta)$ the *continuity modulus generalized* to the space L^p.

Besides $L^p(-\infty, \infty)$, we shall be dealing with L^p spaces in respect to finite intervals for which, in the following, an interval will always be chosen 2π in length. In such case, the functions will be continued beyond this interval pursuant to the relation $f(x + 2\pi) = f(x)$

The generalized modulus of continuity will be denoted in this case by the same symbol as before, but it will now be defined by

$$\omega_p(\delta) = \omega_p(\delta; f) = \sup_{0<h\leq\delta} \left\{ \int_0^{2\pi} |f(x+h) - f(x)|^p \, dx \right\}^{\frac{1}{p}}.$$

Equation (1) holds also for this case and the proof thereof is even simpler.

It will be further observed that in both cases the inequality

$$\omega_p(m\delta) \leqq m\omega_p(\delta)$$

holds, where m is a positive integer.

Whence it follows as above (see Section 79) that the equation

$$\lim_{\delta\to 0} \frac{\omega_p(\delta)}{\delta} = 0$$

entails the identity

$$\omega_p(\delta) \equiv 0 .$$

This identity can hold only in the case in which $f(x)$ is a constant almost everywhere.

In fact, for arbitrary [2] x', x'' $(x' < x'')$ with

$$\frac{1}{p} + \frac{1}{q} = 1$$

we have the inequality

$$\left| \int_x^{x''} \{f(x+h) - f(x)\} \, dx \right| \leqq \{x'' - x'\}^{\frac{1}{p}} \cdot \left\{ \int_{x'}^{x''} |f(x+h) - f(x)|^p \, dx \right\}^{\frac{1}{p}}.$$

Whence there follows

$$\left| \int_{x'}^{x''} \{f(x+h) - f(x)\} \, dx \right| \leqq \{x'' - x'\}^{\frac{1}{q}} \omega_p(h).$$

Hence we may conclude from $\omega_p(\delta) \equiv 0$ that

$$\int_{x'}^{x''} \{f(x+h) - f(x)\} \, dx = 0 ,$$

that is to say

[2] In case we are considering $L^p(0, 2\pi)$, this is of course subject to the condition $x'' - x' \leq 2\pi$.

$$\frac{1}{h}\int_{x''}^{x''+h} f(t)\,dt = \frac{1}{h}\int_{x'}^{x'+h} f(t)\,dt$$

for all values of h, x', x''. If we now choose a fixed x', so that we have

$$\lim_{h\to 0}\frac{1}{h}\int_{x'}^{x'+h} f(t)\,dt = f(x')$$

we find that

$$\lim_{h\to 0}\frac{1}{h}\int_{x'}^{x+h} f(t)\,dt$$

exists for all values of x''. But since this limit is equal to $f(x'')$ for nearly every x'', it follows that for nearly every x'', we have

$$f(x'') = \lim_{h\to 0}\frac{1}{h}\int_{x''}^{x''+h} f(t)\,dt = f(x') = \text{const}\ .$$

When $\omega_p(\delta) \leqq M\delta^\alpha (0 < \alpha \leqq 1)$ we again write $f(x) \in \text{Lip}\,(\alpha, p)$, and we say that $f(x)$ satisfies a Lipschitz condition in the metric L^p.

In place of the generalized modulus of continuity, we may introduce the expression

$$\omega_p^*(\delta) = \omega_p^*(\delta; f) = \sup_{0\leqq h\leqq\delta}\left\{\int\left|f\left(x-\frac{h}{2}\right)+f\left(x+\frac{h}{2}\right)-2f(x)\right|^p dx\right\}^{\frac{1}{p}},$$

which does not exceed $2\omega_p\left(\frac{\delta}{2}\right)$.

Following Zygmund [70] we write $f(x) \in \Lambda_n$ [or $f(x) \in \lambda_p$] when $\omega_p^*(\delta) \leqq$ $\leqq M\delta\left[\text{or } \lim_{\delta\to 0}\frac{\omega_p^*(\delta)}{\delta} = 0\right]$.

81. An Example of Harmonic Approximation [70]. Let us now consider the Weierstrass function

$$f(x) = \sum_{k=1}^{\infty} a^k \cos(b^k x).$$

Here, a is contained in (0,1) and b is an odd number greater than unity (cf. Section 42). Then $f(x)$ is a continuous function and has the period 2π.

Let us investigate the modulus of continuity of this function. For the sake of simplicity, we will take b in the form $b = 4\nu + 1$. When $ab < 1$, $f(x)$ has a continuous derivative. In what follows, it is required that $ab \geqq 1$.

After assigning an arbitrarily chosen number h in (0,1), let us determine a positive integer $n+1$ such that the inequality

$$b^n h \leqq 1 < b^{n+1} h \tag{1}$$

is fulfilled. Let us now write

$$f(x+h) - f(x-h) = -2\sum_{k=1}^{\infty} a^k \sin(b^k h) \sin(b^k x) = P + Q$$

with [3]

$$P = -2\sum_{k=1}^{n} a^k \sin(b^k h) \sin(b^k x),$$

$$Q = -2\sum_{k=n+1}^{\infty} a^k \sin(b^k h) \sin(b^k x).$$

Then it follows that

$$|Q| \leqq 2\sum_{k=n+1}^{\infty} a^k = \frac{2a^{n+1}}{1-a}.$$

On the other hand, however, we have

$$|P| \leqq 2h\sum_{k=1}^{n} a^k b^k,$$

that is

$$|P| \leqq \begin{cases} 2nh & (\text{for } ab = 1), \\ 2abh\dfrac{a^n b^n - 1}{ab-1} < \dfrac{2ba^{n+1}}{ab-1} & (\text{for } ab > 1). \end{cases}$$

Let us now set

$$a = \ln\frac{1}{a} : \ln b, \tag{2}$$

hence, $a = 1$, when $ab = 1$; and otherwise (when $ab > 1$), a lies in (0, 1). Thus we find

$$a^{n+1} < h^a, \quad n \leqq \frac{1}{\ln b} \cdot \ln\frac{1}{h}.$$

Hence, when $ab > 1$, we have

$$|f(x+h) - f(x-h)| < 2\left(\frac{1}{1-a} + \frac{b}{ab-1}\right)h^a.$$

When $ab = 1$, we have

$$|f(x+h) - f(x-h)| < \frac{2}{\ln b} h \cdot \ln\frac{1}{h} + \frac{2h}{1-a}.$$

[3] If $n = 0$, then let $P = 0$.

Accordingly, when $ab > 1$, we get

$$|f(x+h) - f(x)| \leqq Mh^{\alpha}$$

and when $ab \leq 1$,

$$|f(x+h) - f(x)| \leqq M'h \ln \frac{1}{h} \qquad \left(0 < h \leqq \frac{1}{2}\right);$$

here M and M' are independent of both x and h.

Hence, when $ab > 1$, the Weierstrass function satisfies a Lipschitz condition of order α, where α is defined by the formula (2); whereas in the case $ab = 1$ the modulus of continuity $\omega(\delta)$ of this function satisfies the inequality

$$\omega(\delta) \leqq M'\delta \cdot \ln \frac{1}{\delta} .$$

We will now show that our estimates are sharp in the sense that in every interval, no matter how small, there exist two points x', x'' such that in the case $ab > 1$ we have the inequality

$$|f(x'') - f(x')| > m\,|x'' - x'|^{\alpha}$$

and in the case $ab = 1$ we have correspondingly

$$|f(x'') - f(x')| > m'|x'' - x'|\ln \frac{1}{|x'' - x'|} .$$

Here m, m' are positive constants. In the proof we will make use of the assumption not as yet employed, namely, that b has the form $4\nu + 1$.

No matter how small the preassigned interval may be, there always exist therein points of the form

$$\xi = \frac{\pi c}{2b^r}$$

where r is a positive integer and c is an integer of the form $4\mu + 1$. Furthermore, an integer $n > r$ can be chosen so large that both points

$$x' = \xi - h, \quad x'' = \xi + h,$$

wherein we are to put

$$h = \frac{\pi}{b^n} ,$$

still fall within the preassigned interval. Let us assume, besides, that n satisfies the inequality

$$(4 - \pi)\, a^n b^n > (4 + 2\pi)\, a^r b^r$$

when $ab > 1$, and satisfies the inequality

$$(4 - \pi)\, n > (4 + 2\pi)\, r$$

when $ab = 1$.

For the difference $f(x'') - f(x')$ we get

$$f(x'') - f(x') = f(\xi + h) - f(\xi - h) =$$

$$= -2 \sum_{k=1}^{\infty} a^k \sin(b^k \xi) \sin(b^k h) =$$

$$= -2 \sum_{k=1}^{n-1} a^k \sin(b^k \xi) \sin(b^k h),$$

because the equation

$$\sin(b^k h) = \sin(\pi b^{k-n}) = 0$$

is true for $k \geq n$.

If now we set

$$f(x'') - f(x') = R + S$$

with

$$R = -2 \sum_{k=1}^{r-1} a^k \sin(b^k \xi) \sin(b^k h),$$

$$S = -2 \sum_{k=r}^{n-1} a^k \sin(b^k \xi) \sin(b^k h),$$

then obviously

$$S = -2 \sum_{k=r}^{n-1} a^k \sin(b^k h).$$

But since the inequality $\sin x > \frac{2}{\pi} x$ holds for $0 < x < \frac{\pi}{2}$, it follows that

$$|S| > \frac{4}{\pi} \sum_{k=r}^{n-1} h a^k b^k,$$

that is

$$|S| > \begin{cases} \dfrac{4h}{\pi} \dfrac{a^n b^n - a^r b^r}{ab - 1} & (\text{for } ab > 1), \\ \dfrac{4h}{\pi} (n - r) & (\text{for } ab = 1). \end{cases}$$

On the other hand, we have

$$|R| < 2 \sum_{k=1}^{r-1} h a^k b^k,$$

that is

$$|R| < \begin{cases} 2h\dfrac{a^r b^r - ab}{ab-1} & (\text{for } ab > 1), \\ 2h\,(r-1) & (\text{for } ab = 1). \end{cases}$$

If we now recall how n was chosen, it follows from the foregoing that: In the case $ab > 1$, we obtain

$$|f(x'') - f(x')| \geqq |S| - |R| > \\ > \frac{h}{ab-1}\left\{\frac{4}{\pi}(a^n b^n - a^r b^r) - 2\,(a^r b^r - ab)\right\} > \frac{h a^n b^n}{ab-1},$$

i.e.

$$|f(x'') - f(x')| > \frac{\pi^{1-\alpha} h^\alpha}{ab-1} = m\,|x'' - x'|^\alpha;$$

and in the case $ab = 1$, we obtain

$$|f(x'') - f(x')| \geqq |S| - |R| \geqq \frac{4h}{\pi}(n-r) - 2h\,(r-1) > hn = \frac{h \ln \dfrac{\pi}{h}}{\ln b},$$

i.e.

$$|f(x'') - f(x')| > m'\,|x'' - x'| \ln \frac{1}{|x'' - x'|}.$$

This proves our proposition.

Thus we see that the function $f(x)$ does not belong to Lip 1 when $ab = 1$. Considerations wholly analogous to those given above show, however, that in this case the inequality

$$\left|\frac{f(x-h) + f(x+h) - 2f(x)}{h}\right| < \frac{5}{1-\alpha}$$

holds, in other words, that $f(x)$ lies in Λ^*.

Accordingly, a function can satisfy the condition Λ^* without fulfilling a Lipschitz condition of order 1.

82. Some Estimates for Fourier Coefficients. Suppose that $f(x)$ is an integrable function and that

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)\, e^{-ikx}\, dx = \frac{1}{2\pi}\int_a^{2\pi+a} f(x)\, e^{-ikx}\, dx = \\ = \frac{1}{2\pi}\int_0^{2\pi} f(x+a)\, e^{-ik(x+a)}\, dx \qquad (k = 0, \pm 1, \pm 2, \ldots)$$

are its Fourier constants. Let us now assume $k \neq 0$ and let us set $a = \dfrac{\pi}{k}$ in the above formula; then we obtain

$$c_k = -\frac{1}{2\pi}\int_0^{2\pi} f\left(x+\frac{\pi}{k}\right)e^{-ikx}\,dx,$$

$$(1)\qquad c_k = \frac{1}{4\pi}\int_0^{2\pi}\left\{f(x)-f\left(x+\frac{\pi}{|k|}\right)\right\}e^{-ikx}\,dx \qquad (k=\pm 1,\pm 2,\ldots).$$

Hence, if the function $f(x)$ is continuous and if $\omega(\delta)$ is its modulus of continuity, we have

$$|c_k| \leqq \frac{1}{2}\,\omega\left(\frac{\pi}{|k|}\right) \qquad (k=\pm 1,\pm 2,\ldots).$$

If the function $f(x)$ lies in $L^p(p \geqq 1)$ and if $\omega_p(\delta)$ is its generalized modulus of continuity, then, again from (1), we have

$$|c_k| \leqq \frac{1}{4\pi}\,(2\pi)^{\frac{1}{q}}\left\{\int_0^{2\pi}\left|f(x)-f\left(x+\frac{\pi}{k}\right)\right|^p dx\right\}^{\frac{1}{p}} \leqq$$

$$\leqq \frac{1}{2\,(2\pi)^{\frac{1}{p}}}\,\omega_p\left(\frac{\pi}{|k|}\right) \qquad (k=\pm 1,\pm 2,\ldots).$$

We now assume that $f(x)$ has an integrable r-th derivative.

With the aid of the results in Section 52, we then have:

$$|c_k| \leqq \frac{1}{2\,|k|^r}\,\omega^{(r)}\left(\frac{\pi}{|k|}\right) \qquad (k=\pm 1,\pm 2,\ldots),$$

in case the function $f^{(r)}$ is continuous and has $\omega^{(r)}(\delta)$ as its modulus of continuity; and

$$|c_k| \leqq \frac{1}{2\,(2\pi)^{\frac{1}{p}}}\,\frac{1}{|k|^r}\,\omega_p^{(r)}\left(\frac{\pi}{|k|}\right) \qquad (k=\pm 1,\pm 2,\ldots)$$

in case $f^{(r)}$ lies in $L^p(p \geqq 1)$ and has the generalized modulus of continuity $\omega_p^{(r)}(\delta)$.

In some cases the relationship between the differential properties of the function $f(x)$ and the rate of decrease in its Fourier coefficients, which is reflected in the estimates derived above, can be stated in all detail. The simplest theorems pertaining thereto are as follows:

THEOREM 1. *Given,* $f(x) \in L^p(0, 2\pi)$ *and*

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{ikx}.$$

Then

(2) $$f(x) \in \mathrm{Lip}\,(\alpha, 2) \qquad (0 < \alpha < 1)$$

is equivalent to

(3) $$\sum_{|k| \geqq n} |c_k|^2 < \mathrm{const}\,\frac{1}{n^{2\alpha}} \qquad (n \geqq 1)\,.$$

This theorem has a continuous analogue:

THEOREM 2. *Suppose that $f(x)$ is contained in L^2 $(-\infty, \infty)$ and that*

$$f(x) = \underset{\substack{a \to -\infty \\ b \to \infty}}{\mathrm{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_a^b e^{iux}\, \varphi(u)\, du\,,$$

then

$$f(x) \in \mathrm{Lip}\,(\alpha, 2) \qquad (0 < \alpha < 1)$$

is equivalent to

$$\int_{|u| > A} |\varphi(u)|^2\, du < \mathrm{const}\,\frac{1}{A^{2\alpha}} \qquad (A > 0).$$

We will confine ourselves to the proof of the first theorem.
From

$$f\left(x + \frac{\pi h}{4}\right) - f\left(x - \frac{\pi h}{4}\right) \sim 2i \sum_{k=-\infty}^{\infty} c_k\, e^{ikx} \sin\frac{k\pi h}{4}$$

follows

(4) $$\frac{1}{2\pi} \int_0^{2\pi} \left| f\left(x + \frac{\pi h}{4}\right) - f\left(x - \frac{\pi h}{4}\right) \right|^2 dx = 4 \sum_{k=-\infty}^{\infty} |c_k|^2 \sin^2 \frac{k\pi h}{4}.$$

Now, if $f(x)$ lies in Lip $(\alpha, 2)$ $(0 < \alpha < 1)$, then, by reason of the last equation, we have

$$4 \sum_{k=-\infty}^{\infty} |c_k|^2 \sin^2 \frac{k\pi h}{4} \leqq 2Ch^{2\alpha},$$

wherein C is independent of h. Whence it follows that

$$\sum_{\frac{1}{h} \leqq |k| < \frac{2}{h}} |c_k|^2 \leqq Ch^{2\alpha}$$

and accordingly

$$\sum_{\frac{2^r}{h} \leqq |k| < \frac{2^{r+1}}{h}} |c_k|^2 \leqq C\,\frac{h^{2\alpha}}{2^{2r\alpha}} \qquad (r = 0, 1, 2, \ldots)\,.$$

By adding these two inequalities, we find

$$\sum_{k \geqq \frac{1}{h}} |c_k|^2 \leqq C \frac{4^\alpha}{4^\alpha - 1} h^{2\alpha}.$$

Thus (3) follows from (2).

Let us now suppose that (3) is fulfilled. We denote by C_1 the constant appearing in the right-hand side of (3). Then, by reason of (4),

$$\frac{1}{2\pi}\int_0^{2\pi} \left| f\left(x + \frac{\pi h}{4}\right) - f\left(x - \frac{\pi h}{4}\right) \right|^2 dx \leqq 4C_1 h^{2\alpha} + 4 \sum_{k \leqq \frac{1}{h}} |c_k|^2 \frac{\pi^2 h^2}{16} k^2.$$

There only remains

$$I_h = \sum_{k \leqq \frac{1}{h}} h^2 k^2 |c_k|^2$$

to be estimated.

For that purpose, let us set

$$\gamma_k = \sum_{m=k}^{\infty} \{|c_m|^2 + |c_{-m}|^2\} \qquad (k > 0),$$

which gives

$$\gamma_k \leqq C_1 \cdot \frac{1}{k^{2\alpha}} .$$

If we now bring I_h into the form

$$I_h = h^2 \sum_{k=1}^{n} k^2 \{\gamma_k - \gamma_{k+1}\} \qquad \left(n = \left[\frac{1}{h}\right]\right),$$

there follows

$$I_h \leqq h^2 \{\gamma_1 + \gamma_2 (2^2 - 1^2) + \gamma_3 (3^2 - 2^2) + \ldots + \gamma_n (n^2 - [n-1]^2)\} \leqq$$

$$\leqq 2h^2 \sum_{k=1}^{n} k\gamma_k \leqq 2C_1 h^2 \sum_{k=1}^{n} \frac{1}{k^{2\alpha-1}} \leqq C_2 h^2 n^{2-2\alpha} \leqq C_2 h^{2\alpha},$$

where C_2 is independent of h.

Accordingly we have

$$\omega_2\left(\frac{\pi h}{2}\right) \leqq h^\alpha \sqrt{8\pi C_1 + \frac{\pi^3}{2} C_2},$$

and the theorem is proved.

Now, it will be observed that

$$2\pi \sum_{|k| \geqq n} |c_k|^2 = \min_{A_r} \int_0^{2\pi} \left| f(x) - \sum_{r=-n+1}^{n-1} A_r e^{irx} \right|^2 dx .$$

But that means: The left-hand side of the preceding equation is equal to the square of the difference between the best approximation of $f(x)$ through trigonometric sums of order less than the n-th, and $f(x)$ itself, in the metric L^2. Denoting this difference by $\mathcal{T}_n^{(2)}[f]$, we can state Theorem 1 as follows: If

$$\omega_2(\delta; f) \leqq \text{const} \cdot \delta^\alpha \qquad (0 < \alpha < 1),$$

then, for any positive integer n, we have

$$\mathcal{T}_n^{(2)}[f] \leqq \text{const} \cdot \frac{1}{n^\alpha}$$

and conversely.

Theorem 2 can be stated analogously.

Thus we have illustrated to some extent the general considerations of Section 78 in the case of the space L^2.

83. More about V. A. Stekloff's Functions. We are now going to prove some properties of the Stekloff functions (cf. Section 67) wherein the modulus of continuity plays a role.

Suppose that we are given a function $f(x)$ $(-\infty < x < \infty)$ integrable in every finite interval and that its Stekloff function is

$$f_h(x) = \frac{1}{h} \int_{-\frac{h}{2}}^{\frac{h}{2}} f(x+t)\, dt . \tag{1}$$

From this definition, we find the relations

$$f(x) - f_h(x) = \frac{1}{h} \int_{-\frac{h}{2}}^{\frac{h}{2}} \{f(x) - f(x+t)\}\, dt$$

and

$$|f(x) - f_h(x)| \leqq \frac{1}{h} \int_{-\frac{h}{2}}^{\frac{h}{2}} |f(x) - f(x+t)|\, dt. \tag{2}$$

Furthermore — again by reason of (1) — we have for almost every x

$$f'_h(x) = \frac{1}{h}\left\{f\left(x+\frac{h}{2}\right) - f\left(x-\frac{h}{2}\right)\right\}. \tag{3}$$

From (2) and (3) we now find:

a) If $f(x)$ is uniformly convergent along the entire axis, then

$$\sup_{-\infty<x<\infty} |f(x) - f_h(x)| \leqq \omega\left(\frac{h}{2}; f\right),$$

$$\sup_{-\infty<x<\infty} |f'_h(x)| \leqq \frac{1}{h}\,\omega(h; f);$$

b) if $f(x) \in L^p(-\infty, \infty)$ $(p \geqq 1)$ or if $f(x)$ has the period 2π and lies in L^p (π, π) $(p \geqq 1)$, then

$$\|f - f_h\|_p \leqq \omega_p\left(\frac{h}{2}; f\right),$$

$$\|f'_h\|_p \leqq \frac{1}{h}\,\omega_p(h; f).$$

The less trivial of the inequalities given above is

$$\|f - f_h\|_p \leqq \omega_p\left(\frac{h}{2}; f\right).$$

In order to obtain it from (2), we must employ the generalized Minkowski inequality (cf. Section 5).

If we take the mean of the Stekloff function $f_h(x)$, we obtain

$$f_{hh}(x) = \frac{1}{h}\int_{-\frac{h}{2}}^{\frac{h}{2}} f_h(x+t)\,dt.$$

This function has a second derivative

$$f''_{hh}(x) = \frac{1}{h^2}\{f(x+h) + f(x-h) - 2f(x)\}$$

almost everywhere. Besides, it is easy to verify that

$$f(x) - f_{hh}(x) = \frac{1}{h^2}\int_0^{\frac{h}{2}}\int_0^{\frac{h}{2}} \{4f(x) - f(x+s+t) - f(x+s-t) - f(x-s+t) - f(x-s-t)\}\,ds\,dt.$$

With the aid of these relations, we obtain the following estimates:

a*) If $f(x)$ is uniformly continuous along the entire axis, then

$$\sup_{-\infty<x<\infty} |f(x)-f_{hh}(x)| \leqq \frac{1}{2}\omega^*(h;f),$$

$$\sup_{-\infty<x<\infty} |f''_{hh}(x)| \leqq \frac{1}{h^2}\omega^*(2h;f);$$

b*) if $f(x)$ belongs to $L^p(-\infty,\infty)$ $(p \geq 1)$ or if $f(x)$ has the period 2π and lies in $L^p(-\pi,\pi)$ $(p \geq 1)$, then

$$\|f-f_{hh}\|_p \leqq \frac{1}{2}\omega_p(h;f),$$

$$\|f''_{hh}\| \leqq \frac{1}{h^2}\omega^*(2h;f).$$

84. Two Lemmas. We are now going to prove two lemmas which will be useful later.

Lemma 1: *If an integral transcendental function $\Phi(z)$ of the exponential type with an exponent $\leqq T$ has the period $l>0$, then $\Phi(z)$ is a trigonometric sum of the form*

$$\sum_{k=-n}^{n} c_k e^{\frac{2\pi i k}{l}z}$$

with

$$n \leqq \left[\frac{Tl}{2\pi}\right].$$

Proof: Obviously $\Phi(z)$ is bounded along the entire real axis and therefore belongs to B_T. By virtue of Bernstein's inequality, we then have

$$\max_{-\infty<x<\infty} |\Phi^{(m)}(x)| \leqq MT^m \qquad (m=0,1,2,\ldots),$$

wherein

$$M = \max_{-\infty<x<\infty} |\Phi(x)|.$$

Now, let

$$\Phi(x) = \sum_{-\infty}^{\infty} c_k e^{\frac{2\pi i k}{l}x}.$$

Then, for any positive integer m, we have

$$c_k = \frac{1}{l}\left(\frac{l}{2\pi k i}\right)^m \int_0^l \Phi^{(m)}(x)\, e^{-\frac{2\pi i k x}{l}}\, dx$$

and consequently

$$|c_k| \leqq M\left|\frac{Tl}{2\pi k}\right|^m \qquad (m=0,1,2,\ldots).$$

This inequality shows, however, that c_k vanishes when $|k| > \frac{Tl}{2\pi}$.

Lemma 2: If $f(x)$ is a continuous function with the period $l > 0$ and if there exists a function $\Psi(x) \in G_{T-0}$ such that

$$\sup_{-\infty<x<\infty} |f(x) - \Psi(x)| \leqq \delta \,, \tag{1}$$

then there is a trigonometric sum of the form

$$\Phi(x) = \sum_{k=-n}^{n} c_k e^{\frac{2\pi i k x}{l}}$$

with

$$n < \frac{Tl}{2\pi},$$

for which the relation

$$\sup_{-\infty<x<\infty} |f(x) - \Phi(x)| \leqq \delta$$

is likewise fulfilled.

Proof: $f(x)$ is bounded. Then, by reason of (1), $\Psi(z)$, too, is bounded on the real axis. Hence there exists a $\sigma < T$ for which $\Psi(z)$ belongs to B_σ. If we now put

$$\Phi_N(x) = \frac{1}{2N+1} \sum_{k=-N}^{N} \Psi(x + kl),$$

we find, again by reason of (1):

$$\sup_{-\infty<x<\infty} |f(x) - \Phi_N(x)| \leqq \delta \qquad (N = 0, 1, \ldots). \tag{2}$$

On the other hand, the set of functions $\Phi_N(z)$ is bounded uniformly along the real axis. All the functions of the set belong to B_σ. Hence this set is compact (cf. Section 73). Therefore there exists a subsequence $\{\Phi_{N_i}(z)\}$ which converges uniformly in every bounded subdivision of the plane. The limit function $\Phi(z)$ then is likewise contained in B_σ and obviously has the period l. By reason of (2), we have, besides,

$$\sup_{-\infty<x<\infty} |f(x) - \Phi(x)| \leqq \delta.$$

Now we need only apply Lemma 1 and the proof is complete.

85. The Direct Problem of Harmonic Approximation, with reference to the metric of any linear normalized space E, can be formulated for the complex-valued functions defined along the entire real axis, as follows: Given any class of functions $f(x)$ defined along the entire real axis. They need not necessarily belong to the space E. Now, let $\mathfrak{M}_f$ be the set of those functions $F(x)$ in G_{T-0}

for which $f(x) - F(x)$ lies in E and let

$$\varrho_T[f] = \begin{cases} \infty, \text{ if the set } \mathfrak{M}_f \text{ is empty;} \\ \inf\limits_{F \in \mathfrak{M}_f} \| f - F \| \text{ if the set } \mathfrak{M}_f \text{ is not empty.} \end{cases}$$

Required,

$$\varrho_T\{K\} \leqq \sup_{f \in K} \varrho_T[f].$$

Hereafter, we will choose for E on occasion one of the three following spaces:

1° The space of all continuous bounded functions;

2° the space $L^p(-\infty, \infty)$ $(p \geq 1)$;

3° the space $L^p(-\pi, \pi)$ $(p \geq 1)$, whose elements are periodically extended along the entire real axis.

In these cases we will write in place of ϱ_T, the symbols $\mathcal{E}_T$, $\mathcal{J}_T^{(p)}$, $\mathcal{T}_T^{(p)}$ respectively.

The quantity $\mathcal{E}_T\{K\}$ deserves the most attention. It is only a comparatively short time ago that this quantity was determined for a number of important classes K [71], [72]. For other classes, only sporadic estimates of $\mathcal{E}_T\{K\}$ are known. In the spaces L^p we get analogous estimates for the corresponding classes.

The quantity $\mathcal{E}_T\{K\}$ can be determined for classes of functions which admit a representation by means of an integral with a kernel of a particular kind.

We call a function $g(x)$ $(-\infty < x < \infty)$ a *Krein kernel* [72] if

a) $g(x)$ is real-valued and lies in $L(-\infty, +\infty)$;

b) corresponding to every $T > T_0 \geq 0$, there exists a function

$$g_T(x) = \frac{1}{\sqrt{2\pi}} \int_{-T}^{T} e^{itx} \gamma_T(t)\, dt$$

with $\overline{\gamma_T(u)} = \gamma_T(-u)$, $\gamma_T(T) = 0$, $\gamma'_T(u) \in L^2(-T, T)$ such that the product

$$[g(x) - g_T(x)] \sin T(x - a)$$

for certain values of the parameter a, is positive almost everywhere on the real axis;

c) the quantity [4]

$$\int_{-\infty}^{\infty} |g(x) - g_T(x)|\, dx$$

is a continuous function of T.

[4] The existence of this integral follows from the fact that, by virtue of b) we have:

$$(1 + |x|)\, g_T(x) \in L^2(-\infty, \infty).$$

Lemma 1. *If $g(x)$ is a Krein kernel, then we have*

$$\int_{-\infty}^{\infty} |g(x) - g_T(x)|\, dx = \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{1}{2m+1} \Im \{e^{-i\alpha(2m+1)T}\, \zeta([2m+1]T)\} \equiv \Omega_T(\alpha),$$

wherein we are to put

$$\zeta(t) = \int_{-\infty}^{\infty} e^{iux}\, g(x)\, dx\,.$$

Proof: By hypothesis, the product

$$[g(x) - g_T(x)] \sin T(x - \alpha)$$

is positive almost everywhere. Hence we have

$$\int_{-\infty}^{\infty} |g(x) - g_T(x)|\, dx = \int_{-\infty}^{\infty} [g(x) - g_T(x)] \operatorname{sign} \sin T(x - \alpha)\, dx.$$

We now observe: As $n \to \infty$, we find that

$$\frac{4}{\pi} \sum_{k=0}^{n} \frac{\sin (2k+1)\, T(x-\alpha)}{2k+1} \to \operatorname{sign} \sin T(x - \alpha).$$

Here the left-hand side is bounded for all values of n and x (cf. Section 53). Therefore we obtain

$$\int_{-\infty}^{\infty} |g(x) - g_T(x)|\, dx = \lim_{n\to\infty} \frac{4}{\pi} \int_{-\infty}^{\infty} [g(x) - g_T(x)] \sum_{k=0}^{n} \frac{\sin (2k+1)\, T(x-\alpha)}{2k+1}\, dx =$$

$$= \lim_{n\to\infty} \frac{4}{\pi} \int_{-\infty}^{\infty} g(x) \sum_{k=0}^{n} \frac{\sin (2k+1)\, T(x-\alpha)}{2k+1}\, dx\,,$$

since, by virtue of b),

$$\int_{-\infty}^{\infty} g_T(x)\, e^{\pm i(2k+1)Tx}\, dx = 0 \qquad (k = 0, 1, 2, \ldots)\,.$$

Now we need only take into consideration that

$$\int_{-\infty}^{\infty} g(x) \sin u(x-\alpha)\, dx = \Im \int_{-\infty}^{\infty} g(x)\, e^{iu(x-\alpha)}\, dx = \Im \{e^{-iu\alpha}\, \zeta(u)\}\,.$$

Lemma 2. *Let*

$$g_T(x) = \frac{1}{\sqrt{2\pi}} \int_{-T}^{T} e^{iux}\, \gamma_T(u)\, du$$

with $\gamma_T(T) = \gamma_T(-T) = 0$ and $\gamma'_T(u) \in L^2(-T\ T)$. *Suppose, further, that the function* $h(x)$ *satisfies one of the following two conditions*:

a) $h(x) \in W^2$, that is

$$\frac{h(x)}{1+|x|} \in L^2(-\infty, \infty),$$

b) $h(x) \in W^1$, that is

$$\frac{h(x)}{1+|x|} \in L^1(-\infty, \infty).$$

Then the function

$$\varphi(x) = \int_{-\infty}^{\infty} g_T(x-u)\, h(u)\, du = \int_{-\infty}^{\infty} g_T(u)\, h(x-u)\, du$$

lies in G_T *as well as in* W^2.

Proof: If we set

$$h_1(x) = \frac{h(x)}{x-i},$$

then $\varphi(x)$ takes the form

$$\varphi(x) = (x-i) \int_{-\infty}^{\infty} g_T(x-u)\, h_1(u)\, du - \int_{-\infty}^{\infty} (x-u)\, g_T(x-u)\, h_1(u)\, du.$$

a) Let $h_1(x)$ be contained in $L^2(-\infty, \infty)$, so that we have

$$h_1(x) = \operatorname*{l.i.m.}_{\substack{a \to -\infty \\ b \to \infty}} \frac{1}{\sqrt{2\pi}} \int_a^b e^{ixt}\, \Theta(t)\, dt$$

with $\Theta(t) \in L^2(-\infty, \infty)$. Since, by hypothesis, both $g_T(x)$ and $xg_T(x)$ belong to $L^2(-\infty, +\infty)$, it follows from the theorem on the Faltung of two functions in $L^2(-\infty, \infty)$, that

$$\varphi(x) = (x-i) \int_{-T}^{T} e^{ixt}\, \Theta(t)\, \gamma_T(t)\, dt - i \int_{-T}^{T} e^{ixt}\, \Theta(t)\, \gamma'_T(t)\, dt.$$

Thus Lemma 2 has been proved for the case a).

b) Suppose, now, that $h(x)$ belongs to $L^1(-\infty, \infty)$. Let us set

$$\varphi_1(x) = \int_{-\infty}^{\infty} g_T(x-u)\, h_1(u)\, du,$$

$$\varphi_2(x) = \int_{-\infty}^{\infty} (x-u)\, g_T(x-u)\, h_1(u)\, du,$$

so that we have

$$\varphi(x) = (x-i)\, \varphi_1(x) - \varphi_2(x).$$

By virtue of Fubini's theorem, $\varphi_1(x)$ lies in $L^1(-\infty, +\infty)$. We will now prove that each of the two functions $\varphi_1(x)$, $\varphi_2(x)$ coincides with a function lying in B_T and restricted to the real axis. Thus the assertion in Lemma 2 will also be established for the case b).

The function $zg_T(z)$ lies in B_T. By virtue of Bernstein's inequality, (see Section 74), the relation

$$\sup_{-\infty<x<\infty} \left| \frac{d^r}{dx^r} [x g_T(x)] \right| \leqq M T^r \qquad (r = 1, 2, \ldots)$$

with

$$M = \sup_{-\infty<x<\infty} |x g_T(x)|$$

is then fulfilled.

Consequently,

$$\sup_{-\infty<x<\infty} |\varphi_2^{(r)}(x)| \leqq M T^r \int_{-\infty}^{\infty} |h_1(u)|\, du \qquad (r = 0, 1, 2, \ldots),$$

whence (see Section 73) we can conclude that $\varphi_2(x)$ is identical with a function $\varphi_2(z)$ lying in B_T and restricted to the real axis. Just so, it can be proved also that $\varphi_1(z)$ lies in B_T.

It will be noted that at least one of the conditions a), b), is satisfied if $h(x)$ is bounded or if it belongs to $L^p(-\infty, \infty)$ $(p \geqq 1)$.

In the first instance, the function $\varphi(x)$ is also bounded; in the second, $\varphi(x)$ lies also in $L^p(-\infty, \infty)$ $(p \geqq 1)$, which follows from Minkowski's inequality.

Furthermore, it is clear that $\varphi(x)$ becomes a trigonometric sum if $h(x)$ possesses a period. In such case, it is no longer necessary to assume that $h(x)$ belongs to W^2. The theorem of Hardy-Young (see Section 63) allows this proposition provided that $h(x)$ lies in $L(-\pi, \pi)$. In fact,

$$g'_T(x) = \frac{1}{\sqrt{2\pi}} \int_{-T}^{T} e^{iux}\, iu \gamma_T(u)\, du,$$

whence there follows

$$(1+|x|)\, g_T'(x) \in L^2(-\infty, \infty)\ .$$

Accordingly we have

$$g_T'(x) \in L^1(-\infty, \infty).$$

Hence $g_T(x)$ is of bounded variation along the entire real axis.

THEOREM 1. *Let K be the totality of all functions $f(x)$ $(-\infty < x < \infty)$ of the form*

$$f(x) = \int_{-\infty}^{\infty} g(x-t)\, h(t)\, dt + F_c(x), \tag{1}$$

where $h(t)$ runs through the set of all measurable functions for which

$$\operatorname*{true\,max}_{-\infty<t<\infty} |h(t)| \leqq 1\ .$$

Here we suppose that $g(x)$ is a Krein kernel (which characterizes the class K) and that $F_c(x)$ is a function dependent on $f(x)$ and lying in G_τ. Here c is independent of $f(x)$.

Then for $T > \max\{T_0, c\}$ we have

$$\mathcal{E}_T\{K\} = \Omega_T(x) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1}\, \Im\{e^{-i\alpha(2k+1)T}\, \zeta([2k+1]T)\}.$$

Proof: Take any c which satisfies the inequality

$$\max\{T_0, c\} < \tau < T$$

and construct the function

$$\varphi(x) = \int_{-\infty}^{\infty} g_\tau(x-t)\, h(t)\, dt.$$

By virtue of Lemma 2, the function $\varphi(x)$ (which is bounded along the entire real axis) lies in $G_\tau \subset G_{T-0}$. Therefore

$$\Phi(x) = \varphi(x) + F_c(x)$$

also lies in G_{T-0}. On the other hand,

$$f(x) - \Phi(x) = \int_{-\infty}^{\infty} \{g(x-t) - g_\tau(x-t)\}\, h(t)\, dt$$

and accordingly,

$$\sup_{-\infty<x<\infty} |f(x) - \Phi(x)| \leqq \int_{-\infty}^{\infty} |g(x) - g_\tau(x)|\, dx.$$

Hence it follows that

$$\mathcal{E}_T[f] \leqq \int_{-\infty}^{\infty} |g(x) - g_\tau(x)|\, dx.$$

We now allow τ to approach T and utilize only the property c) of Krein's kernel. We then obtain

$$\mathcal{E}_T[f] \leqq \Omega_T(\alpha)$$

and

$$\mathcal{E}_T\{K\} \leqq \Omega_T(\alpha).$$

In order to prove that the equality sign holds in this relation, let us consider the function

$$f_0(x) = \int_{-\infty}^{\infty} g(x-t)\,\operatorname{sign}\sin T(x_0 - t)\, dt =$$

$$= \int_{-\infty}^{\infty} \{g(x-t) - g_T(x-t)\}\operatorname{sign}\sin T(x_0 - t)\, dt,$$

where x_0 is chosen arbitrarily. By virtue of the calculations made above, we then have

$$f_0(x) = \Omega_T(x - x_0)$$

and

$$\max_{-\infty < x < \infty} |f_0(x)| = \Omega_T(\alpha).$$

Now, since $f_0(x)$ has the period $\frac{2\pi}{T}$, it follows from Lemma 2 of Section 84 that the quantity $\mathcal{E}_T[f_0]$ is equal to the difference between $f_0(x)$ and the best approximation thereof through trigonometric sums

$$\sum_{k=-n}^{n} c_k e^{ikTx}$$

of the order $n < 1$, that is through constants. According to Tchebysheff's theorem for periodic functions (Section 40), these constants are equal to zero. Hence the function $\Omega_T(x - x_0)$ assumes its maximum value $\Omega_T(\alpha)$, with opposite signs, at the two points

$$x = x_0 + \alpha, \quad x = x_0 + \alpha + \frac{\pi}{T}$$

in the interval of length $\frac{2\pi}{T}$. Therefore we have

$$\mathcal{E}_T[f_0] = \Omega_T(\alpha).$$

But since $f_0(x)$ belongs to K, the theorem is proved.

THEOREM 2. *If $f(x)$ has the form* (1), *if $g(x)$ is a Krein kernel, if $h(x) \in L^p$ $(-\infty, \infty)$ $(p \geq 1)$ and $F_c(x)$ belongs to G_c, then we have*

$$\mathscr{J}_T^{(p)}[f] \leqq \Omega_T(a) \, \| h \|_p$$

when $T > \max\{T_0, c\}$.

Proof: By Lemma 2 and the observations made thereon, the function

$$\varphi(x) = \int_{-\infty}^{\infty} g_\tau(u) \, h(x-u) \, du$$

lies in G_τ. The same thing is true of

$$\Phi(x) = \varphi(x) + F_c(x).$$

On the other hand, we have

$$f(x) - \Phi(x) = \int_{-\infty}^{\infty} \{g(u) - g_\tau(u)\} \, h(x-u) \, du,$$

and the generalized Minkowski inequality leads to

$$\| f - \Phi \|_p \leqq \| h \|_p \int_{-\infty}^{\infty} | g(x) - g_\tau(x) | \, dx.$$

Whence, by the same reasoning as that in the proof of Theorem 1, there follows the inequality

$$\mathscr{J}_T^{(p)}[f] \leqq \Omega_T(a) \, \| h \|_p.$$

In an analogous manner, we can prove

THEOREM 3. *If a function $f(x)$ with the period 2π has the form* (1), *if $g(x)$ is a Krein kernel, $h(x)$ a periodic function in $L^p(-\pi, \pi)$ $(p \geq 1)$ and $F_c(x)$ a trigonometric sum of the order $\leqq c$, then, for $T > \max\{T_0, c\}$ we have*

$$\mathscr{J}_T^{(p)}[f] \leqq \Omega_T(a) \, \| h \|_p.$$

In conclusion, we wish to call attention to a certain analogy between the considerations of this section and those of Section 50, where we discussed the researches of A. A. Markoff.

86. A Criterion Due to B. Sz. Nagy [73]. If we wish to apply the results of the preceding section to this or that concrete class K which admits a representation of the form

$$f(x) = \int_{-\infty}^{\infty} g(x-t) \, h(t) \, dt + F_c(x)$$

we must verify beforehand that the conditions a), b), c) imposed on the function $g(x)$ are actually fulfilled. Of these conditions, b) is fundamental and specific. Frequently, the following criterion due to Nagy enables us to determine whether condition b) is fulfilled:

Let $g(x)$ $(-\infty < x < \infty)$ *be an*

odd | *even*

continuous function in $L(-\infty, \infty)$. *Furthermore, suppose that the function*

$$\lambda(x) = \int_{-\infty}^{\infty} g(t) \sin tx\, dt \qquad | \qquad \lambda(x) = \int_{-\infty}^{\infty} g(t) \cos tx\, dt$$

has a derivative $\lambda'(x) \in L^2(-T, T)$ *in the interval* $[-T, T]$ $(T > 0)$ *and that* $\lambda(x)$ *is*

two-fold | *three-fold*

monotonic in the interval $[T, \infty]$. *Besides, let* $\lambda(T) \neq 0$. *Then the function*

$$g_T(x) = \frac{1}{\pi} \int_0^T \{\lambda(t) - \mu(t)\} \sin tx\, dt \qquad | \qquad g_T(x) = \frac{1}{\pi} \int_0^T \{\lambda(t) - \mu(t)\} \cos tx\, dt,$$

with

$$\mu(t) = \sum_{k=1}^{\infty} \{\lambda(2kT - t) - \lambda(2kT + t)\}; \qquad | \qquad \mu(t) = \sum_{k=1}^{\infty} (-1)^{k-1} \{\lambda(2kT - t) + \lambda(2kT + t)\}$$

belongs to $L(-\infty, \infty)$, *and the difference*

$$g(x) - g_T(x)$$

changes sign at exactly the same points as

$\sin Tx$ | $\cos Tx$.

Let us first assume that $g(x)$ is odd. Obviously, the function defined by the equation

$$\mu(t) = \sum_{k=1}^{\infty} \{\lambda(2kT - t) - \lambda(2kT + t)\} \qquad (0 \leqq t \leqq T)$$

is monotonic and fulfills the equations

$$\mu(0) = 0, \qquad \mu(T) = \lambda(T).$$

Since $\lambda(x)$ decreases monotonically in the interval $[T, \infty]$ and converges to zero as x approaches ∞, the integral

$$\int_T^{\infty} \lambda'(x)\, dx$$

converges. Accordingly the series

$$\sum_{k=1}^{\infty} \{\lambda'(2kT - t) + \lambda'(2kT + t)\}$$

converges uniformly in $[0, T]$. Hence $\mu(t)$ has a continuous derivative. Consequently we have

$$xg_T(x) = \frac{1}{\pi} \int_0^T \{\lambda'(t) - \mu'(t)\} \cos xt \, dt.$$

Whence there follows $xg_T(x) \in L^2(-\infty, \infty)$, that is $g_T(x) \in L(-\infty, \infty)$.

Let us now consider the difference

$$g(x) - g_T(x) = \frac{1}{\pi} \int_T^{\to\infty} \lambda(t) \sin xt \, dt + \frac{1}{\pi} \int_0^T \sum_{k=1}^{\infty} \{\lambda(2kT - t) - \lambda(2kT + t)\} \sin xt \, dt.$$

Since

$$\int_0^T \sum_{k=N+1}^{\infty} \{\lambda(2kT - t) - \lambda(2kT + t)\} \sin xt \, dt$$

approaches zero as N goes to ∞ and since, by virtue of the existence of

$$\int_A^{\to\infty} \lambda(t) \sin xt \, dt, \qquad \int_A^{\to\infty} \lambda(t) \cos xt \, dt$$

the expression

$$\int_T^{\to\infty} \lambda(2NT + t) \sin xt \, dt =$$

$$= \cos 2NTx \int_{(2N+1)T}^{\to\infty} \lambda(t) \sin xt \, dt - \sin 2NTx \int_{(2N+1)T}^{\to\infty} \lambda(t) \cos xt \, dt$$

has the same property, we obtain

$$g(x) - g_T(x) = \lim_{N \to \infty} I_N$$

with

$$I_N = \frac{1}{\pi} \left\{ \int_T^{\to\infty} [\lambda(t) - \lambda(2NT + t)] \sin xt \, dt + \right.$$

$$\left. + \int_0^T \sum_{k=1}^{N} [\lambda(2kT - t) - \lambda(2kT + t)] \sin xt \, dt \right\}.$$

Let us now set

$$\sigma_N(t) = \sum_{k=0}^{N-1} \lambda(t + [2k+1]T).$$

Then a simple rearrangement gives

$$I_N = \frac{1}{\pi} \int_0^{\to\infty} \{\sigma_N(|T-v|) - \sigma_N(T+v)\} \sin xv\, dv.$$

Whence there follows

$$I_N = 2 \sin Tx \cdot \frac{1}{\pi} \int_0^{\to\infty} \sigma_N(u) \cos xu\, du = \sin Tx \cdot S_N(x),$$

wherein we are to put

$$S_N(x) = \frac{2}{\pi} \int_0^{\to\infty} \sigma_N(u) \cos xu\, du.$$

Since $\sigma_N(u)$ is doubly monotonic in the interval $[0, \infty]$, it follows that $S_N(x)$ is non-negative for any value of x and an arbitrary N (cf. Section 68). Therefore the sign of

$$g(x) - g_T(x)$$

is the same as that of $\sin Tx$ for all values of x, as we were to prove.

Now we will consider the case where $g(x)$ is even. The function

$$\mu(t) = \sum_{k=1}^{\infty} (-1)^{k-1} \{\lambda(2kT - t) + \lambda(2kT + t)\}$$

can be brought into the form

$$\mu(t) = \varrho(T + t) + \varrho(T - t)$$

wherein

$$\varrho(v) = \sum_{k=1}^{\infty} (-1)^{k-1} \lambda(v + [2k-1]T).$$

By hypothesis, $\varrho(v)$ is a doubly monotonic function on $[0, \infty]$ for which

$$\varrho(0) > 0, \quad \varrho(v) \to 0 \qquad (v \to \infty).$$

Now, we observe that $\mu(T) = \lambda(T)$ and find:

$$xg_T(x) = \frac{1}{\pi} \int_0^T \{\mu'(t) - \lambda'(t)\} \sin xt\, dt.$$

Consequently, $xg_T(x)$ lies in $L^2(-\infty, \infty)$, that is, $g_T(x)$ lies in $L(-\infty, \infty)$.

Now let us consider the difference

$$g(x) - g_T(x).$$

When $x \neq 0$, we can bring this difference into the form

$$g(x) - g_T(x) = \frac{1}{\pi}\int_T^{\to\infty} \lambda(t)\cos xt\,dt + \frac{1}{\pi}\int_{-T}^{T} \varrho(t+T)\cos xt\,dt\,.$$

Furthermore, when $t \geq T$, we have

$$\varrho(t-T) + \varrho(t+T) = \lambda(t).$$

It can be shown without difficulty that, for $x \neq 0$, the following holds:

$$g(x) - g_T(x) = \frac{1}{\pi}\int_T^{\to\infty} \{\varrho(t-T) + \varrho(t+T)\}\cos xt\,dt +$$

$$+ \frac{1}{\pi}\int_{-T}^{T} \varrho(t+T)\cos xt\,dt = 2\cos Tx \cdot \frac{1}{\pi}\int_0^{\to\infty} \varrho(u)\cos xu\,du.$$

Now we need only recall once more the theorem in Section 68.

In concluding this section it should be observed also that in both cases under consideration the function $\Omega_T(\alpha)$ of Section 85 is equal to

$$\frac{4}{\pi}\sum_{k=0}^{\infty} \frac{\lambda([2k+1]T)}{2k+1}$$

when $g(x)$ is odd, and is equal to

$$\frac{4}{\pi}\sum_{k=0}^{\infty} (-1)^k \frac{\lambda([2k+1]T)}{2k+1}$$

when $g(x)$ is even.

87. The Best Approximation of Differentiable Functions. In 1911 D. Jackson [66] proved the following important theorem:

Given: *$f(x)$ is a continuous function with the period 2π, has an r-th derivative $(r > 1)$, and satisfies the inequality*

$$|f^{(r)}(x)| \leqq 1\,.$$

To prove: *There exists for any positive integer n a trigonometric sum $\mathcal{S}_{n-1}(x)$ of order less that n; this trigonometric sum satisfies the inequality*

$$|f(x) - \mathcal{S}_{n-1}(x)| \leqq \frac{A_r}{n^r} \qquad (0 \leqq x \leqq 2\pi),$$

and hence the value of A_r depends solely on the value of r.

We will designate this theorem as the first Jackson theorem. It answers one of a number of questions contained in the general problem which was formulated in Section 78. A comparatively short time ago (cf. the first two papers under [71]), a new, simple proof of Jackson's first theorem was discovered; this proof enables us to find the best value of the quantity A_r.

This proof, which emerges almost without any preliminary considerations, will be given later. In the present Section, on the other hand, we will show that Jackson's theorem and certain of its generalizations are special cases of the general theorem in Section 85.

Let us now take an arbitrary positive number c and an even function $p(t)$ which can satisfy the following conditions:

1. $p(t) = 1$, for $t \geq c$ and for $t \leq -c$;
2. $p(t)$ has a continuous second derivative;
3. The ratios

$$\frac{p(t)}{t^{r+2}}, \quad \frac{p'(t)}{t^{r+1}}, \quad \frac{p''(t)}{t^r}$$

do not become infinite.

Furthermore, let us set

$$\lambda(t) = \frac{p(t)}{(it)^r}$$

and

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} \lambda(t)\, dt,$$

wherein $g(x)$ is defined as

$$\lim_{\omega \to \infty} \frac{1}{2\pi} \int_{-\omega}^{\omega} e^{itx}\, \lambda(t)\, dt$$

when $r = 1$. Accordingly $g(x)$ is a real function which is even for even values and odd for odd values of r.

An integration by parts shows at once that

$$(1 + x^2)\, g(x) \in L^2(-\infty, \infty).$$

Whence it follows, among other things, that

$$(1 + |x|)\, g(x) \in L(-\infty, \infty).$$

Let us now introduce the conjugate function $\tilde{g}(x)$, which obviously belongs to $L(-\infty, \infty)$ and is odd for even values of r, and even for odd values of r. With the aid of Nagy's criterion, we can conclude from the foregoing that the two functions $g(x)$, $\tilde{g}(x)$ are Krein kernels. Simple calculations, which may be omitted here, lead to the relations

$$\int_{-\infty}^{\infty} |g(x) - g_T(x)|\, dx = \frac{4}{\pi T^r} \sum_{k=0}^{\infty} \frac{(-1)^{k(r+1)}}{(2k+1)^{r+1}}, \tag{1}$$

$$\int_{-\infty}^{\infty} |\tilde{g}(x) - \tilde{g}_T(x)|\, dx = \frac{4}{\pi T^r} \sum_{k=0}^{\infty} \frac{(-1)^{kr}}{(2k+1)^{r+1}}. \tag{2}$$

We will now assume that $f(x)$ $(-\infty < x < \infty)$ has an r-th derivative

$$f^{(r)}(x) = h(x),$$

which belongs to W^2 and show that then we have

$$f(x) = \int_{-\infty}^{\infty} g(x-t)\, h(t)\, dt + F_c(x) \tag{3}$$

with $F_c(x) \in G_c$.

To that end, let us set

$$h(t) = h_1(t)\,(t-i)$$

and let $\psi_1(t)$ denote the Fourier transform of $h_1(t)$. Since

$$I(x) \equiv \int_{-\infty}^{\infty} g(x-s)\, h(s)\, ds =$$

$$= (x-i) \int_{-\infty}^{\infty} g(x-s)\, h_1(s)\, ds - \int_{-\infty}^{\infty} (x-s)\, g(x-s)\, h_1(s)\, ds$$

it follows by reason of the theorem on the Faltung of two functions in $L^2(-\infty, \infty)$ that

$$I(x) = (x-i) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx}\, \lambda(t)\, \psi_1(t)\, dt - \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx}\, \lambda'(t)\, \psi_1(t)\, dt. \tag{4}$$

From the appearance of equation (4) it is evident that $I(x)$ has a derivative of order $r-1$ which is absolutely continuous in every finite interval. This derivative is equal to

$$I^{(r-1)}(x) = (x-i) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (it)^{r-1}\, \lambda(t)\, \psi_1(t)\, e^{itx}\, dt +$$

$$+ (r-1) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (it)^{r-2}\, \lambda(t)\, \psi_1(t)\, e^{itx}\, dt - \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (it)^{r-1}\, \lambda'(t)\, \psi_1(t)\, e^{itx}\, dt.$$

We are now going to prove that $I^{(r-1)}(x)$ has the derivative

$$(5) \qquad I^{(r)}(x) = (x-i)\frac{1}{\sqrt{2\pi}}\,\text{l. i. m.}_{\substack{a\to-\infty\\ b\to\infty}}\int_a^b (it)^r \lambda(t)\,\psi_1(t)\,e^{itx}\,dt +$$

$$+ r\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(it)^{r-1}\lambda(t)\,\psi_1(t)\,e^{itx}\,dt - \frac{i}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(it)^r\lambda'(t)\,\psi_1(t)\,e^{itx}\,dt$$

almost everywhere. In fact,

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(it)^{r-1}\lambda(t)\,\psi_1(t)\,e^{itx}\,dt = q(x)$$

is continuous and satisfies the relation

$$q(x) - q(0) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{e^{itx}-1}{it}(it)^r\lambda(t)\,\psi_1(t)\,dt.$$

The function

$$(it)^r\lambda(t)\,\psi_1(t)$$

lies in $L^2(-\infty,\infty)$. Therefore, by Plancherel's theorem (cf. Section 60),

$$q(x) - q(0)$$

has the derivative

$$q'(x) = \frac{1}{\sqrt{2\pi}}\,\text{l. i. m.}_{\substack{a\to-\infty\\ b\to\infty}}\int_a^b e^{itx}(it)^r\lambda(t)\,\psi_1(t)\,dt$$

almost everywhere.

Now the proof of formula (5) presents no further difficulty.

This formula can be written in the following form:

$$I^r(x) = h(x) - \frac{i}{\sqrt{2\pi}}\int_{-c}^{c}e^{itx}p'(t)\,\psi_1(t)\,dt + (x-i)\frac{1}{\sqrt{2\pi}}\int_{-c}^{c}e^{itx}[pt)(-1]\,\psi_1(t)\,dt$$

$$= h(x) - \Phi_c(x) = f^{(r)}(x) - \Phi_c(x)$$

with $\Phi_c(x)\in G_c$. But since

$$\frac{d^r}{dx^r}\{f(x) - I(x)\} = \Phi_c(x)$$

it follows that $f(x) - I(x)$ is equal to the r-tuple integral of $\Phi_c(x)$, that is to say $f(x) - I(x)$ is likewise a function $F_c(x)$ in G_c. Thus formula (3) has been established.

The right-hand side of formula (4) changes sign if and only if $\lambda(t)$, $\psi_1(t)$ are equal to $i \operatorname{sign} t \cdot \lambda(t)$, $i \operatorname{sign} t \cdot \psi_1(t)$, respectively. Hence, by virtue of the definition of the conjugate function, we have

$$I(x) = -\int_{-\infty}^{\infty} \tilde{g}(x-t)\,\tilde{h}(t)\,dt.$$

Consequently, the formula

$$f(x) = -\int_{-\infty}^{\infty} \tilde{g}(x-t)\,\tilde{h}(t)\,dt + F_c(x) \tag{6}$$

is likewise valid, jointly with (3). From the general theorem in Section 85 and the formulas (3), (6) derived above, we obtain

THEOREM 1. *Let K_r, $\tilde{K}_r (r \geq 1)$, be the totalities of all functions $f(x)$ $(-\infty < x < \infty)$ for which we have*

$$\operatorname*{true\,max}_{-\infty < x < \infty} |h(x)| \leq 1,$$

$$\operatorname*{true\,max}_{-\infty < x < \infty} |\tilde{h}(x)| \leq 1,$$

respectively, wherein we are to put $h(x) = f^{(r)}(x)$. Then, for any positive value of T,

$$\mathscr{E}_T\{K_r\} = \frac{\mathsf{K}_r}{T^r},$$

$$\mathscr{E}_T\{\tilde{K}_r\} = \frac{\tilde{\mathsf{K}}_r}{T^r},$$

respectively, with

$$\mathsf{K}_r = \frac{4}{\pi}\sum_{m=0}^{\infty} \frac{(-1)^{m(r+1)}}{(2m+1)^{r+1}},$$

$$\tilde{\mathsf{K}}_r = \frac{4}{\pi}\sum_{m=0}^{\infty} \frac{(-1)^{mr}}{(2m+1)^{r+1}}$$

respectively.

The theorem of Jackson formulated at the beginning of this Section, as well as the precise value of the constant A_r, are merely special cases of Theorem 1. It will be observed that the equation

$$\mathscr{E}_T\{K_r\} = \frac{\mathsf{K}_r}{T^r}$$

remains valid when $r = 0$ ($\mathbf{K}_0 = 1$) if

$$|f(x)| \leqq 1$$

holds, not only almost everywhere, but everywhere.

Moreover, the following estimates are valid:

$$1 \leqq \mathbf{K}_0 < \mathbf{K}_2 < \mathbf{K}_4 < \ldots < \frac{4}{\pi} < \ldots < \mathbf{K}_5 < \mathbf{K}_3 < \mathbf{K}_1 = \frac{\pi}{2}$$

and

$$\tilde{\mathbf{K}}_1 < \tilde{\mathbf{K}}_3 < \tilde{\mathbf{K}}_5 < \ldots < \frac{4}{\pi} < \ldots < \tilde{\mathbf{K}}_6 < \tilde{\mathbf{K}}_4 < \bar{\tilde{\mathbf{K}}}_2 < \frac{\pi}{2}.$$

Accordingly, $\frac{\pi}{2}$ is the common least upper bound of all the constants given in Theorem 1.

We will now pass to the approximation of the differentiable functions in L^p and prove first that the formula (3) is valid for almost every x if the function $h(x) = f^{(r)}(x)$ satisfies the following two conditions:

a) for arbitrary finite values of α, β ($\alpha < \beta$) we have

$$\int_\alpha^\beta |h(x)|\,dx \leqq M + N(\beta - \alpha),$$

where M, N are independent of α, β.

b) $$\int_0^x h(t)\,dt = f^{(r-1)}(x) - f^{(r-1)}(0) \in W^2.$$

Both conditions are obviously satisfied if $h(x) \in L^p(-\pi, \pi)$ and has a period; a) and b) are likewise satisfied [5] if $h(x)$ lies in $L^p(-\infty, \infty)$, where $1 \leqq p < 2$. In fact, if $h(x)$ is periodic, $f^{(r-1)}(x)$ is bounded. If, on the other hand, $h(x)$ lies in $L^p(-\infty, \infty)$ ($1 \leqq p < 2$), then, by Hölder's inequality, we have

$$\int_\alpha^\beta |h(x)|\,dx \leqq (\beta - \alpha)^{\frac{1}{q}} \|h\|_p \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right),$$

$$|f^{(r-1)}(x) - f^{(r-1)}(0)| \leqq |x|^{\frac{1}{q}} \|h\|_p.$$

[5] If $f(x)$ lies in $L^p(-\infty, \infty)$ ($p \geqq 2$), then $f(x)$ belongs automatically also to W^2. For $p = 2$, the statement is obviously true; for $p > 2$, it follows from Hölder's inequality, that is

$$\int_{-\infty}^{\infty} \frac{|h(x)|^2}{(1 + |x|)^2}\,dx \leqq \left\{\int_{-\infty}^{\infty} |h(x)|^p\,dx\right\}^{\frac{2}{p}} \left\{\int_{-\infty}^{\infty} \frac{dx}{(1 + |x|)^s}\right\}^{1 - \frac{2}{p}}$$

with $s = \frac{2p}{p-2} > 2$. Hence, in the case where $p \geqq 2$, the formulas (3), (6) are unnecessary for the proof.

This proves our proposition, since $q > 2$ by hypothesis.

Suppose the conditions a) and b) are satisfied. Let us consider the function

$$\int_0^x f(t)\,dt,$$

which has an $(r-1)$-st derivative

$$f^{(r-1)}(x)$$

belonging to W^2. For this function the formula (3) is valid. Accordingly,

$$\int_0^x f(t)\,dt = \int_{-\infty}^{\infty} g(t)\,[f^{(r-1)}(x-t) - f^{(r-1)}(-t)]\,dt + F_c(x).$$

Therefore

$$\int_0^x f(t)\,dt = \int_{-\infty}^{\infty} g(t)\,dt \int_0^x h(s-t)\,ds + F_c(x).$$

From

$$\int_{-\infty}^{\infty} |g(t)|\,dt \int_0^{|x|} |h(s-t)|\,dt \leqq \{M + N|x|\} \int_{-\infty}^{\infty} |g(t)|\,dt < \infty$$

it follows from Fubini's theorem that

$$\int_0^x f(t)\,dt = \int_0^x ds \int_{-\infty}^{\infty} g(t)\,h(s-t)\,dt + F_c(x).$$

According to Lebesgue's theorem, we then have almost everywhere

$$f(x) = \int_{-\infty}^{\infty} g(t)\,h(x-t)\,dt + F_c'(x),$$

which was to be proved.

The formula (6) admits a similar generalization. As before, let

$$\varphi(x) \equiv f^{(r-1)}(x) - f^{(r-1)}(0) \in W^2.$$

Let us assume that the function $\tilde{\varphi}(x)$, the conjugate of $\varphi(x)$, has the form

$$\tilde{\varphi}(x) = \int_0^x \psi(t)\,dt,$$

where

$$\int_{\alpha}^{\beta} |\psi(t)|\,dt \leqq M + N(\beta - \alpha)$$

for arbitrary finite values of α, β ($\beta > \alpha$). As a generalization of (6), we then have almost everywhere

$$f(x) = -\int_{-\infty}^{\infty} \tilde{g}(t)\,\psi(x-t)\,dt + F_c'(x).$$

If $h(x) = f^{(r)}(x) \in W^2$, which is certainly true when $h(x) \in L^p(-\infty, \infty)$ [6] ($p \geqq 2$), then $\psi(x)$ is conjugate to $h(x)$ in the sense of Section 69. We can now formulate the theorems of approximation arising from Theorems 2, 3 of Section 85.

THEOREM 2. *From*

$$\int_{-\infty}^{\infty} |f^{(r)}(x)|^p\,dx \leqq 1 \qquad (p \geqq 1,\ r \geqq 0)$$

the inequality

$$\mathscr{S}_T^{(p)}[f] \leqq \frac{\mathsf{K}_r}{T^r}$$

follows for an arbitrary positive T.
If the function

$$\int_{0}^{x} \psi(t)\,dt$$

is the conjugate of $f^{(r-1)}(x)$ ($r \geqq 1$), and if

$$\int_{-\infty}^{\infty} |\psi(t)|^p\,dx \leqq 1 \qquad (p \geqq 1),$$

then, for an arbitrary positive T, it follows that

$$\mathscr{S}_T^{(p)}[f] \leqq \frac{\tilde{\mathsf{K}}_r}{T^r}.$$

THEOREM 3. *If a periodic function $f(x)$ satisfies the condition*

$$\int_{-\pi}^{\pi} |f^{(r)}(x)|^p\,dx \leqq 1 \qquad (p \geqq 1),$$

[6] Cf. footnote [5].

then, for an arbitrary positive T, we have

$$\mathscr{S}_T^{(p)}[f] \leqq \frac{\mathsf{K}_r}{T^r} \qquad (r = 0, 1, 2, \ldots)$$

and

$$\mathscr{S}_T^{(p)}[\tilde{f}] \leqq \frac{\tilde{\mathsf{K}}_r}{T^r} \qquad (r = 1, 2, 3, \ldots).$$

88. Direct Observations Concerning Periodic Functions. We begin with the calculation of the expressions

$$\mathsf{K}_{r,n} = \min_{A_m} \frac{1}{\pi} \int_{-\pi}^{\pi} \left| \varphi_r(t) - \sum_{-(n-1)}^{n-1} A_m e^{imt} \right| dt,$$

$$\tilde{\mathsf{K}}_{r,n} = \min_{A_m} \frac{1}{\pi} \int_{-\pi}^{\pi} \left| \psi_r(t) - \sum_{-(n-1)}^{n-1} A_m e^{imt} \right| dt,$$

where $\varphi_r(t)$, $\psi_r(t)$ are the functions introduced in Section 55, and $r \geqq 1$. Since $\varphi_r(t)$ and $\psi_r(t)$ are even and odd respectively, it is easy to see that we have

$$\mathsf{K}_{r,n} = \min_{\xi_m} \frac{2}{\pi} \int_0^{\pi} \left| \varphi_r(t) - \sum_{m=0}^{n-1} \xi_m \cos\left(mt - \frac{\pi r}{2}\right) \right| dt,$$

$$\tilde{\mathsf{K}}_{r,n} = \min_{\eta_m} \frac{2}{\pi} \int_0^{\pi} \left| \psi_r(t) - \sum_{m=0}^{n-1} \eta_m \sin\left(mt - \frac{\pi r}{2}\right) \right| dt,$$

where ξ_m, η_m run through real values.

We will utilize the explanations given in Section 51 and prove that neither of the expressions

$$\Phi_r(t) = \varphi_r(t) - \sum_{m=0}^{n-1} \xi_m \cos\left(mt - \frac{\pi r}{2}\right),$$

$$\Psi_r(t) = \psi_r(t) - \sum_{=0}^{n-1} \eta_m \sin\left(mt - \frac{\pi r}{2}\right)$$

for abitrary real values of ξ_m, η_m can change sign in the interval $0 < t < \pi$ more often than the functions $\cos\left(nt - \frac{\pi r}{2}\right)$, $\sin\left(nt - \frac{\pi r}{2}\right)$ respectively. Then, by virtue of Section 51, it would follow that

$$\mathsf{K}_{r,n} = \frac{2}{\pi}\left|\int_0^{\pi} \varphi_r(t)\,\operatorname{sign}\cos\left(nt - \frac{\pi r}{2}\right)dt\right|,$$

$$\tilde{\mathsf{K}}_{r,n} = \frac{2}{\pi}\left|\int_0^{\pi} \psi_r(t)\,\operatorname{sign}\sin\left(nt - \frac{\pi r}{2}\right)dt\right|.$$

By evaluating these integrals, we get

$$\mathsf{K}_{r,n} = \frac{\mathsf{K}_r}{n^r}, \qquad \tilde{\mathsf{K}}_{r,n} = \frac{\tilde{\mathsf{K}}_r}{n^r},$$

where K_r, $\tilde{\mathsf{K}}_r$ are the quantities defined in Section 87.

Without loss of generality, we may let r be even. Let us now assume that the number of zeros counted properly [7] of $\Phi_r(t)$ in the interval $0 < t < \pi$ is equal to $N > n$ (for specially selected constants ξ_m).

According to Rolle's theorem, the function

$$\Phi_{r-1}(t) - \Phi_r'(t)$$

has at least $N - 1$ zeros in the interval $0 < t < \pi$. Moreover, $t = 0, \pi$ are zeros. Whence it follows — again by Rolle's theorem — that $\Phi_{r-2}(t)$ also has not less than N zeros in the interval $0 < t < \pi$. By repeating this reasoning, we arrive at the result that the function

$$\Phi_1(t) = \frac{1}{2}t - \sum_{m=1}^{n-1} m^{r-1}\,\xi_m \sin mt$$

has at least $N - 1$ zeros in the interval $0 < t < \pi$. But since $\Phi_1(t)$, too, vanishes for $t = 0$, it follows that

$$\frac{1}{2} - \sum_{m=1}^{n-1} m^r\,\xi_m \cos mt$$

has at least $N - 1 > n - 1$ zeros in the interval $0 < t < \pi$. But that is evidently absurd. This proves the proposition for $\Phi_r(t)$.

We now consider $\Psi_r(t)$. Let N, the number of zeros of $\Psi_r(t)$, be greater than $n - 1$ in the interval $0 < t < \pi$ for properly chosen constants η_m. Since $\Psi_r(t)$ vanishes also for $t = 0$ and $t = \pi$, the function

$$\Psi_{r-2}(t) = \Psi_r''(t),$$

according to Rolle's theorem, again has at least N zeros in the interval $0 < t < \pi$. By repeating this reasoning, we find that the function

[7] That is, with due care as to multiple zeros.

$$\Psi_2(t) = \psi_2(t) + \sum_{m=1}^{n-1} m^{r-2}\eta_m \sin mt$$

has at least n zeros in the interval $0 < t < \pi$ and that

$$\Psi_1(t) = \ln\left(2\cos\frac{t}{2}\right) + \sum_{m=1}^{n-1} m^{r-1}\eta_m \cos mt$$

has at least $N + 1$ zeros in the same interval. Therefore the function

$$\frac{1}{2}\operatorname{tg}\frac{t}{2} + \sum_{m=1}^{n-1} m^r\eta_m \sin mt$$

has no less than $N > n - 1$ zeros in the interval $0 < t < \pi$. The same thing holds true for the function

$$1 + 2\sum_{m=1}^{n-1} m^r\eta_m \cos\frac{t}{2}\frac{\sin mt}{\sin\frac{t}{2}} = \sum_{m=0}^{n-1}\zeta_m \cos^m t \not\equiv 0$$

(because $\cos\frac{t}{2}\frac{\sin mt}{\sin\frac{t}{2}}$ is a polynomial of the m-th degree in $\cos t$). But that is absurd. Hence our proposition has also been established for the function $\Psi_r(t)$.

Hereafter, it will be understood that ξ_m, η_m mean such values of the coefficients already denoted above by the same symbols, for which the expressions considered by us assume their minimum.

Let $f(x)$ be a function possessing an integrable r-th derivative ($r \geq 1$); in addition, let

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

$$\tilde{f}(x) = \sum_{k=-\infty}^{\infty} i \operatorname{sign} k \cdot c_k e^{ikx}.$$

Without loss of generality we may take $c_0 = 0$.

Let us now construct the trigonometric sums

$$\mathscr{X}_{n-1}(x) = \mathscr{X}_{n-1}(x; f; r) = \sum_{m=1}^{n-1} (-1)^{m-1} m^r \xi_m [c_m e^{imx} + c_{-m} e^{-imx}],$$

$$\mathscr{Y}_{n-1}(x) = \mathscr{Y}_{n-1}(x; f; r) = \sum_{m=1}^{n-1} (-1)^m m^r \eta_m i [c_m e^{imx} - c_{-m} e^{-imx}].$$

By means of simple calculations it is easy to verify the following two equations

$$(1)\quad \begin{cases} \mathscr{L}_{n-1}(x) - f(x) = \\ \quad = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} \left\{ \varphi_r(t) - \sum_{m=0}^{n-1} \xi_m \cos\left(mt - \frac{\pi r}{2}\right) \right\} f^{(r)}(x + \pi - t)\, dt, \\ \mathscr{T}_{n-1}(x) - \tilde{f}(x) = \\ \quad = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} \left\{ \psi_r(t) - \sum_{m=0}^{n-1} \eta_m \sin\left(mt - \frac{\pi r}{2}\right) \right\} f^{(r)}(x + \pi - t)\, dt. \end{cases}$$

Then, if

$$(2)\qquad \text{true max } |f^{(r)}(x)| \leqq 1,$$

it follows by virtue of the minimum-problem [8] solved above that

$$(3)\qquad |f(x) - \mathscr{L}_{n-1}(x)| \leqq \frac{K_r}{n^r}; \qquad |\tilde{f}(x) - \mathscr{T}_{n-1}(x)| \leqq \frac{\tilde{K}_r}{n^r}.$$

From the first of these inequalities we get Jackson's first theorem.

The estimates thus obtained cannot be further sharpened. This can be shown in essentially the same way as in Section 85, through the investigation of a special function.

For example, in the case of the first inequality, let us choose the function

$$f_0(x) = \frac{4}{\pi n^r} \sum_{k=0}^{\infty} \frac{\sin\left([2k+1]\, nx - \frac{\pi r}{2}\right)}{(2k+1)^{r+1}}.$$

For this function we have

$$f_0^{(r)}(x) = \operatorname{sign} \sin nx\,.$$

$f_0(x)$ assumes the values

$$x_\lambda = \frac{\lambda\pi}{n} + \frac{\pi\varepsilon}{2n} \qquad (\lambda = 0, 1, \ldots, 2n-1)$$

with alternate change in sign, at the consecutive points $\frac{K_r}{n^r}$ of an interval of length 2π. Here we may put ε equal to 1 if r is even, and equal to zero if r is odd.

If we consider $f_0(x)$ as the difference between $f_0(x)$ and that trigonometric

[8] If the inequality

$$|f(x)| \leqq 1$$

holds for all values of x, then the first estimate in (3) is satisfied also for $r = 0$.

sum of order $n-1$ which is identically zero, we find, by virtue of de la Vallée-Poussin's theorem on periodic functions (cf. Section 40), that the error in the best approximation of $f_0(x)$ through trigonometric sums below the n-th order cannot be less than $\frac{\mathsf{K}_r}{n^r}$. But that is equivalent to saying the error is $\frac{\mathsf{K}_r}{n^r}$.

In our hypothesis, let us replace (2) by $f^{(r)}(x) \in L^p(-\pi, \pi)$ $(p \geqq 1)$ and let $\| f^{(r)} \|_p \leqq 1$. Then, from the relations (1), it follows that

$$\| f - \mathscr{X}_{n-1} \|_p \leqq \frac{\mathsf{K}_r}{n^r}; \qquad \| \tilde{f} - \mathscr{Y}_{n-1} \|_p \leqq \frac{\tilde{\mathsf{K}}_r}{n^r}.$$

The first of these inequalities holds for $r = 0$ as well.

Thus, in the case of periodic functions, we have derived a new proof of the results in the preceding Section.

It will be observed that the sums $\mathscr{X}_{n-1}(x)$, $\mathscr{Y}_{n-1}(x)$ will be identically zero if every

$$c_k = 0 \qquad (k = 0, \pm 1, \pm 2, \ldots, \pm (n-1)).$$

Whence we have the result: *From*

$$f(x) = \sum_{|k| \geqq n} c_k e^{ikx}$$

and

$$|f^{(r)}(x)| \leqq 1 \qquad (r \geqq 1)$$

it follows that

$$|f(x)| \leqq \frac{\mathsf{K}_r}{n^r}, \qquad |\tilde{f}(x)| \leqq \frac{\tilde{\mathsf{K}}_r}{n^r}.$$

These inequalities cannot be sharpened.

In order to construct the trigonometric sums $\mathscr{X}_{n-1}(x)$, $\mathscr{Y}_{n-1}(x)$ it is necessary to know the constants ξ_m, η_m; these constants in turn depend on r and n. It is not difficult to derive explicit expressions for ξ_m, η_m; however, we will not go into this matter.

89. Jackson's Second Theorem. Let $f(x)$ $(-\infty < x < \infty)$ be a function which is r-fold differentiable $(r \geqq 0)$. Suppose $f(x)$ is integrable in every finite interval. Setting $h = \frac{2}{T}$ $(T > 0)$, let us construct the Stekloff function $f_h(x)$ and bring $f(x)$ into the form

$$f(x) = \{f(x) - f_h(x)\} + f_h(x) = \Phi_h(x) + f_h(x).$$

If $f^{(r)}(x)$ is uniformly continuous on the entire number axis, then, for all values of x (cf. Section 83), we have

$$|\Phi_h^{(r)}(x)| \leqq \omega\left(\frac{1}{T}; f^{(r)}\right),$$

$$|f_h^{(r+1)}(x)| \leqq \frac{T}{2}\,\omega\left(\frac{2}{T}; f^{(r)}\right) \leqq T\,\omega\left(\frac{1}{T}; f^{(r)}\right).$$

Whence, by virtue of Theorem 1, Section 87, we obtain, for an arbitrary positive value of T,

$$\mathscr{E}_T[f] \leqq \frac{K_r}{T^r}\,\omega\left(\frac{1}{T}; f^{(r)}\right) + \frac{K_{r+1}}{T^{r+1}}\,T\omega\left(\frac{1}{T}; f^{(r)}\right) \leqq \frac{3}{T^r}\,\omega\left(\frac{1}{T}; f^{(r)}\right).$$

If, moreover, $f^{(r)}(x)$ has a uniformly continuous conjugate $\tilde{f}^{(r)}(x)$, we have, for an arbitrary positive value of T and $r \geqq 1$,

$$\mathscr{E}_T[f] \leqq \frac{3}{T^r}\,\omega\left(\frac{1}{T}; \tilde{f}^{(r)}\right).$$

In the special case where $f(x)$ has the period 2π, it follows from the foregoing, by Lemma 2 in Section 84, that, for any positive integer n, there exists a trigonometric sum $\mathscr{S}_{n-1}(x) = \mathscr{S}_{n-1}(x; f; r)$ of order $n-1$ for which

$$\max_{0 \leqq x \leqq 2\pi} |f(x) - \mathscr{S}_{n-1}(x)| \leqq \frac{3}{n^r}\,\omega\left(\frac{1}{n}; f^{(r)}\right). \tag{1}$$

If $r \geqq 1$, there exists a second trigonometric sum $\mathscr{T}_{n-1}(x) = \mathscr{T}_{n-1}(x; f; r)$ for which the inequality

$$\max_{0 \leqq x \leqq 2\pi} |\tilde{f}(x) - \mathscr{T}_{n-1}(x)| \leqq \frac{3}{n^r}\,\omega\left(\frac{1}{n}; f^{(r)}\right) \tag{2}$$

is satisfied.

By employing the notation of Section 88, we can write, for example:

$$\mathscr{S}_{n-1}(x) = \mathscr{X}_{n-1}(x; \Phi_h; r) + \mathscr{X}_{n-1}(x; f_h; r+1),$$
$$\mathscr{T}_{n-1}(x) = \mathscr{Y}_{n-1}(x; \Phi_h; r) + \mathscr{Y}_{n-1}(x; f_h; r+1)$$

with

$$\mathscr{X}_{n-1}(x; \Phi; 0) \equiv 0.$$

The inequality (1) expresses the second Jackson theorem [66], [75].

It is easy to see that analogous formulas hold for the approximation of periodic functions in $L^p(-\pi, \pi)$:

$$\mathscr{S}_n[f] \leqq \|f - \mathscr{S}_{n-1}\|_p \leqq \frac{3}{n^r}\,\omega_p\left(\frac{1}{n}; f^{(r)}\right) \qquad (r = 0, 1, 2, \ldots),$$

$$\mathscr{S}_n[\tilde{f}] \leqq \|\tilde{f} - \mathscr{T}_{n-1}\|_p \leqq \frac{3}{n^r}\,\omega_p\left(\frac{1}{n}; f^{(r)}\right) \qquad (r = 1.2, 3, \ldots).$$

It will be further observed that, by Theorem 2 of Section 87, the inequality

$$\mathscr{E}_T^{(p)}[f] \leqq \frac{3}{T^r}\,\omega_p\left(\frac{1}{T}; f^{(r)}\right) \qquad (r = 0, 1, 2, \ldots)$$

holds for arbitrary positive values of T in case $f^{(r)}(x)$ belongs to $L^p(-\infty, \infty)$. Moreover, it will be noted that

$$\mathscr{E}_T^{(p)}[f] \leqq \frac{3}{T^r}\,\omega_p\left(\frac{1}{T};\psi\right)$$

if $f^{(r-1)}(x)$ $(r \geq 1)$ has the conjugate

$$\int_0^x \psi(t)\,dt$$

where $\psi(t) \in L^p(-\infty, \infty)$.

In concluding this Section, let us observe further that, instead of Stekloff's function $f_h(x)$, the function $f_{hh}(x)$ can also serve. The latter function has been introduced in Section 83. From the identity

$$f(x) = \{f(x) - f_{hh}(x)\} + f_{hh}(x)$$

and the inequalities a) of Section 83, we obtain the following result: If $f(x)$ possesses an r-th derivative $(r = 0, 1, 2, \ldots)$, then

$$\mathscr{E}_T[f] \leqq \frac{\pi}{2T^r}\,\omega^*\left(\frac{3}{T};f^{(r)}\right),$$

for all positive values of T. Finally,

$$\mathscr{E}_T[f] \leqq \frac{\pi}{2T^r}\,\omega^*\left(\frac{3}{T};\tilde{f}^{(r)}\right)$$

with $r \geq 1$, if $f^{(r)}(x)$ has a conjugate $\tilde{f}^{(r)}(x)$.

Analogous estimates hold for the approximation in the metric L^p.

90. The Generalized Fejér Method. From the proofs carried out above, in establishing the theorems pertaining to the approximation of differentiable functions, there emerge certain methods for the actual construction of approximations. However, these methods have a shortcoming: They require the calculation of quantities which depend on r. We will now show that, for the classes of differentiable functions studied hereinabove, there exist methods of approximation which are free of this shortcoming and which furnish an approximation of the same order as that furnished by the functions considered above. But now the numerical coefficients in the expressions for the least upper bound of the deviation turn out to be larger. Such a method is attained, for example, by means of a certain modification [76] of Fejér's method.

Let us choose any number $\Theta(0 < \Theta \leq 1)$ and consider the integral

$$\sigma_T(x;f;\Theta) = \int_0^\infty \left\{f\left(x + \frac{u}{\Theta T}\right) + f\left(x - \frac{u}{\Theta T}\right)\right\} \Phi(u)\,du \tag{1}$$

with

$$\Phi(u) = \Phi(u;\Theta) = \frac{\cos\frac{1-\Theta}{\Theta}u - \cos\frac{u}{\Theta}}{\pi u^2}.$$

The function $\Phi(u)$ is a kernel of the Fejér type and reduces to the classical Fejér kernel when $\Theta = 1$ (cf. Section 61).

With a view of obtaining additional estimates, let us prove:

$$\lambda(\Theta) = 2\int_0^\infty |\Phi(u;\Theta)|\,du \leqq \frac{4}{\pi} + \frac{2}{\pi}\ln\frac{2-\Theta}{\Theta}. \tag{2}$$

In fact, we have

$$\lambda(\Theta) = \frac{4}{\pi}\int_0^\infty \left|\frac{\sin\frac{(2-\Theta)u}{2\Theta}\sin\frac{u}{2}}{u^2}\right| du =$$

$$= \frac{4}{\pi}\left\{\int_0^{\frac{2\Theta}{2-\Theta}} + \int_{\frac{2\Theta}{2-\Theta}}^{2} + \int_2^\infty \left|\frac{\sin\frac{(2-\Theta)u}{2\Theta}\sin\frac{u}{2}}{u^2}\right| du\right\}.$$

If we now replace the quantity

$$\left|\sin\frac{(2-\Theta)u}{2\Theta}\sin\frac{u}{2}\right|$$

in the first integral by $\frac{2-\Theta}{2\Theta}u^2$, in the second by $\frac{u}{2}$ and in the third by 1, we obtain the inequality

$$\lambda(\Theta) \leqq \frac{4}{\pi}\frac{2\Theta}{2-\Theta}\frac{2-\Theta}{4\Theta} + \frac{2}{\pi}\log\frac{2-\Theta}{\Theta} + \frac{2}{\pi} = \frac{4}{\pi} + \frac{2}{\pi}\ln\frac{2-\Theta}{\Theta}.$$

The kernel $\Phi(u)$ can be written in the following form:

$$\Phi(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi(\Theta v)\cos uv\,dv$$

with

$$\varphi(t) = 2\int_0^\infty \Phi(u)\cos\frac{ut}{\Theta}\,du = \begin{cases} 1 & \text{for } |t| \leqq 1-\Theta, \\ \frac{1}{\Theta}(1-|t|) & \text{for } 1-\Theta \leqq |t| \leqq 1, \\ 0 & \text{for } |t| \geqq 1. \end{cases} \tag{3}$$

Let us now turn to the generalized Fejér integral (1) and enumerate some of its properties.

1°. If $f(x)$ belongs to W^2 or to W^1, then $\sigma_T(x; f; \Theta) \in G_T$. This is a consequence of Lemma 2 of Section 85.

2°. If $f(x)$ belongs to G_T when $\tau \leq T(1-\Theta)$ and if $f(x)$ lies in W^2, then

$$\sigma_T(x; f; \Theta) \equiv f(x).$$

In this case, as we know,

$$f_1(x) = \frac{f(x) - f(i)}{x - i} = \int_{-\tau}^{\tau} e^{itx} \psi(t)\, dt$$

with $\psi(t) \in L^2$. It will now be observed that

$$\sigma_T(x; f; \Theta) = \Theta T \int_{-\infty}^{\infty} f(x-v)\, \Phi(\Theta T v)\, dv =$$

$$= f(i) + (x-i) \int_{-\infty}^{\infty} f_1(x-v) \frac{\cos T(1-\Theta)v - \cos Tv}{\pi \Theta T v^2}\, dv -$$

$$- \int_{-\infty}^{\infty} f_1(x-v) \frac{\cos T(1-\Theta)v - \cos Tv}{\pi \Theta T v}\, dv\ ;$$

hence, by the theorem on the Faltung of functions, we have

$$\sigma_T(x; f; \Theta) = f(i) + (x-i) \int_{-\tau}^{\tau} \varphi\left(\frac{t}{T}\right) \psi(t)\, e^{itx}\, dt - \frac{i}{T} \int_{-\tau}^{\tau} \varphi'\left(\frac{t}{T}\right) \psi(t)\, e^{itx}\, dt.$$

Now we need merely to recall (3), and the proof is complete.

3°. If $f(x)$ is an integrable function with the period 2π and if

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{ikx},$$

then it follows that

$$\sigma_T(x; f; \Theta) = \sum_{|k| < T} \varphi\left(\frac{k}{T}\right) c_k e^{ikx}.$$

This is a consequence of the general statements made in Section 64.

4°. If n is a positive integer, then, under the hypothesis in 3°, we have:

$$\sigma_n\left(x; f; \frac{1}{n}\right) = s_{n-1}(x)$$

is the Fourier sum of order $n-1$ of $f(x)$.

We now pass to the evaluation of the rapidity with which differentiable functions are approximated by means of the generalized Fejér integral, and assume that $f(x)$ belongs to W^2. Relative to our hypothesis imposed on $f^{(r)}(x)$, this does not mean any loss of generality. In fact, we are putting into effect the representation (3) of Section 87. This representation shows that every function under consideration is the sum of two functions, one belonging to W^2, the other, to G_c (c being positive and as small as we please). When $T > c$, it follows that the approximation is based solely on the first term of this sum.

The reasoning in Section 85 shows that in dealing with the approximation of functions belonging to W^2, we can limit ourselves to those approximating functions which likewise belong to W^2.

Let $f(x)$ be a function belonging to W^2 and having an r-th derivative $f^{(r)}(x)$ which satisfies the hypothesis of any theorem in Section 87. Then, corresponding to the arbitrary positive numbers ε and r, there exists a function $F(x) \in G_{\tau-0}$ which likewise belongs to W^2 and which satisfies the inequality

$$\|f - F\| \leqq \varrho_\tau[f] + \varepsilon \ .$$

Let us take a positive number $\Theta < 1$ and set

$$\tau = T(1 - \Theta) .$$

By virtue of property 2° of the generalized Fejér integral, we then have

$$f(x) - \sigma_T(x; f; \Theta) = f(x) - F(x) - \sigma_T(x; f - F; \Theta) .$$

Whence the inequality

$$\|f(x) - \sigma_T(x; f; \Theta)\| \leqq \|f - F\| + \|\sigma_T(x; f - F; \Theta)\|$$

follows. It will now be observed that

$$\|\sigma_T(x; f - F; \Theta)\| \leqq \lambda(\Theta)\, \|f - F\| \ ;$$

hence we obtain

$$\|f(x) - \sigma_T(x; f; \Theta)\| \leqq [1 + \lambda(\Theta)]\, \varrho_\tau[f] + \varepsilon\, [1 + \lambda(\Theta)] .$$

But since ε is any positive number, it follows that

$$\|f(x) - \sigma_T(x; f; \Theta)\| \leqq [1 + \lambda(\Theta)]\, \varrho_\tau[f] . \tag{4}$$

We treat in the same way the case in which one of the following two inequalities is fulfilled:

$$\varrho_\tau[f] \leqq \frac{A}{\tau^r} ,$$

$$\varrho_\tau[f] \leqq \frac{\Omega\left(\frac{1}{\tau}\right)}{\tau^r} ,$$

where $\Omega(\delta)$ is a non-decreasing function which satisfies the inequality

$$\Omega(\mu\delta) \leqq (\mu + 1)\, \Omega(\delta)$$

for an arbitrary positive μ. From (4) we get for these two cases

$$\| f(x) - \sigma_T(x; f; \Theta) \| \leqq \frac{1 + \lambda(\Theta)}{(1 - \Theta)^r} \frac{A}{T^r}$$

$$\| f(x) - \sigma_T(x; f; \Theta) \| \leqq \frac{2[1 + \lambda(\Theta)]}{(1 - \Theta)^{r+1}} \frac{\Omega\left(\frac{1}{T}\right)}{T^r},$$

respectively.

By applying the generalized Fejér integral, we thus obtain a rapidity in approximating of the same order as before, except that now the numerical coefficients in the expressions for the rapidity are somewhat larger.

In concluding this Section, we call attention to some additional facts which connect with our discussion and relate to the rapidity in the approximating (in the metric of the space C) of continuous periodic functions by means of their Fourier sums $s_{n-1}(x)$.

a) If

$$|f(x)| \leqq M,$$

almost everywhere,

then it follows that $\max |s_{n-1}(x)| \leqq (A + B \ln n)\, M$,

where A, B are constants and $A < 2$, $B < 1$. This is a theorem due to Lebesgue. For the proof, it suffices to consider property 4° of the generalized Fejér integral and the inequality

$$\lambda\left(\frac{1}{n}\right) \leqq \frac{4}{\pi} + \frac{2}{\pi} \ln (2n - 1).$$

b) If $f(x)$ has an r-th derivative $f^{(r)}(x)$ $(r \geqq 0)$ with the property

$$|f^{(r)}(x)| \leqq M_r,$$

then we have [9]

$$\max |f(x) - s_{n-1}(x)| \leqq \frac{A + 1 + B \ln n}{(n-1)^r} M_r.$$

c) If $f(x)$ has an r-th derivative $f^{(r)}(x)$ $(r \geqq 1)$ with the modulus of continuity $\omega^{(r)}(\delta)$, then we have

$$\max |f(x) - s_{n-1}(x)| \leqq \frac{6\,(A + 1 + B \ln n)}{(n-1)^r} \omega^{(r)}\left(\frac{1}{n}\right).$$

[9] A new and interesting result due to A. M. Kolmogoroff [77] attaches to this old and well-known inequality.

d) If $r \geq 1$, the estimates in b) and c) are valid also for

$$\max |\tilde{f}(x) - \tilde{s}_{n-1}(x)|.$$

When $r = 0$, it follows from c) that the Fourier series of a function $f(x)$ converges uniformly if $f(x)$ satisfies the Dini-Lipschitz condition (cf. Section 79).

91. Bernstein's Theorems [65] are converses of Jackson's theorems. These converse theorems answer questions which are typical of those raised in the "converse problems" of harmonic approximation.

THEOREM 1. *If, corresponding to every $T \geq T_0 > 0$, there exists in G_T a function $F_T(x)$ which satisfies the inequality*

$$\sup_{-\infty<x<\infty} |f(x) - F_T(x)| \leqq \frac{A}{T^{r+\alpha}} \tag{1}$$

where A, the integer $r \geq 0$ and α $(0 < \alpha < 1)$ are all three independent of T, then we have:

$$f(x) = F_{T_0}(x) + \varphi(x),$$

wherein $\varphi(x)$ is a function bounded along the entire axis and having a bounded derivative of the r-th order belonging to Lip α.

Proof: Let us set

$$T_k = 2^k T_0 \qquad (k = 0, 1, 2, \ldots)$$

and consider the series

$$F_{T_0}(x) + \{F_{T_1}(x) - F_{T_0}(x)\} + \{F_{T_2}(x) - F_{T_1}(x)\} + \ldots \equiv$$
$$\equiv F_{T_0}(x) + \Phi_1(x) + \Phi_2(x) + \ldots,$$

which obviously converges to $f(x)$ uniformly along the entire real axis.

Since $\Phi_k(x) = F_{T_k}(x) - F_{T_{k-1}}(x) \in G_{T_k}$ $(k = 1, 2, \ldots)$ and because, by reason of (1),

$$\sup_{-\infty<x<\infty} |\Phi_k(x)| \leqq \sup_{-\infty<x<\infty} |f(x) - F_{T_k}(x)| + \sup_{-\infty<x<\infty} |f(x) - F_{T_{k-1}}(x)| \leqq$$
$$\leqq \frac{A}{T_k^{r+\alpha}} + \frac{A}{T_{k-1}^{r+\alpha}} = \frac{B}{T_k^{r+\alpha}},$$

it follows from Bernstein's inequality (cf. Section 74) that

$$\sup_{-\infty<x=\infty} \left| \frac{d^r}{dx^r} \Phi_k(x) \right| \leqq \frac{B}{T_k^{r+\alpha}} T_k^r = \frac{B}{T_k^\alpha} \qquad (k = 1, 2, \ldots).$$

Therefore the series

$$\sum_{k=1}^{\infty} |\Phi_k(x)|, \qquad \sum_{k=1}^{\infty} |\Phi_k^{(r)}(x)|$$

converge uniformly along the entire axis. But since $\Phi_k(x)$ belongs to B_{T_k} and is uniformly continuous along the entire axis, the bounded function

$$\varphi(x) = \sum_{k=1}^{\infty} \Phi_k(x)$$

has the bounded derivative of the r-th order

$$\varphi^{(r)}(x) = \sum_{k=1}^{\infty} \Phi_k^{(r)}(x),$$

which is uniformly continuous along the entire axis.

We now preassign a sufficiently small but otherwise arbitrary positive number δ, then determine a positive integer m from the condition

$$2^{m-1} \leqq \frac{1}{\delta} < 2^m$$

and set

$$\varphi^{(r)}(x) = \left(\sum_{k=1}^{m-1} + \sum_{k=m}^{\infty} \right) \Phi_k^{(r)}(x) \equiv P(x) + Q(x).$$

Then we have the estimate

$$|Q(x)| \leqq \frac{B}{T_0^a} \sum_{k=m}^{\infty} \frac{1}{2^{ka}} = \frac{2^a B}{T_0^a (2^a - 1)} \frac{1}{2^{ma}} < C_1 \delta^a.$$

On the other hand, for $|h| \leq \delta$ — again by reason of Bernstein's inequality — we have the estimate

$$|P(x+h) - P(x)| \leqq \delta \sum_{k=1}^{m-1} \sup_{-\infty < x < \infty} |\Phi_k^{(r+1)}(x)| \leqq$$

$$\leqq B\delta T_0^{1-a} \sum_{k=1}^{m-1} 2^{k(1-a)} < \frac{B\delta T_0^{1-a} 2^{m(1-a)}}{2^{1-a} - 1} < C_2 \delta^a.$$

Accordingly, for $|h| \leq \delta$ we obtain the inequality

$$|\varphi^{(r)}(x+h) - \varphi^{(r)}(x)| < (2C_1 + C_2)\, \delta^a,$$

which proves the theorem.

From the estimates attained in the foregoing proof, we have for $a = 1$:

$$|Q(x+h) - Q(x)| < 2C_1\delta,$$

$$|P(x+h) - P(x)| \leqq B\delta(m-1) \leqq \frac{B}{\ln 2}\, \delta \cdot \ln \frac{1}{\delta}.$$

Whence we obtain

THEOREM 2. *If, corresponding to every* $T \geq T_0 > 0$ *in* G_T, *there exists a function* $F_T(x)$ *satisfying the inequality*

$$\sup_{-\infty < x < \infty} |f(x) - F_T(x)| \leqq \frac{A}{T^{r+1}}, \tag{2}$$

wherein A *and the integer* $r \geq 0$ *do not depend on* T, *then there exists a representation of the form*

$$f(x) = F_{T_0}(x) + \varphi(x).$$

Here $\varphi(x)$ *is a function which is bounded along the entire axis and has a bounded* r*-th derivative whose modulus of continuity is*

$$\omega(\delta; \varphi^{(r)}) \leqq M\delta \ln \frac{1}{\delta}.$$

It will be noted that, when $r \geq 1$, not only $\varphi^{(r)}(x)$ but also the function conjugate to $\varphi^{(r)}(x)$ is bounded. That statement applies, of course, to both theorems and is the result of employing dominant functions analogous to the foregoing with the aid of the inequality which constitutes the substance of the second theorem in Section 74.

Theorem 1 is a complete converse of the corresponding theorem of Jackson. Things are quite different with Theorem 2, because now the factor $\ln \frac{1}{\delta}$ enters in the estimate furnished by means of (2). We will show by an example that Theorem 2 cannot be sharpened.

Let us consider the periodic case and take the Weierstrass function

$$f(x) = \sum_{k=1}^{\infty} a^k \cos(b^k x),$$

where b is an integer greater than 1 and is of the form $4\nu + 1$. Let $a = \frac{1}{b}$. We already know (Section 81) that this function has

$$M\delta \ln \frac{1}{\delta}$$

for its modulus of continuity.

On the other hand, it was established in Section 42 that, for an arbitrary value of n, the error in the best approximation of $f(x)$ through trigonometric sums of the n-th order is

$$\frac{1}{1-a} \cdot \frac{1}{n}.$$

Consequently, in the case where $a = 1$, the logarithmic factor belongs to the essence of the matter and is not the result of too rough an approximation.

In order to arrive at a determinate result for $\alpha = 1$, it is necessary to pass from Lip 1 to Λ^*, as Zygmund [70] was the first to observe recently.

THEOREM 3. *Under the hypothesis of Theorem* 2, *we have*

$$f(x) = F_{T_0}(x) + \varphi(x).$$

Here $\varphi(x)$ is a function bounded along the entire axis and having a bounded derivative of the r-th order belonging to Λ^*.

In the beginning, the proof does not differ from the proof of Theorem 1. After deriving the relation

$$\varphi^{(r)}(x) = P(x) + Q(x)$$

and the inequality

$$|Q(x)| < C_1\delta$$

let us consider

$$\Delta^2 P = P\left(x + \frac{h}{2}\right) + P\left(x - \frac{h}{2}\right) - 2P(x) =$$

$$= \sum_{k=1}^{m-1} \left\{ \Phi_k^{(r)}\left(x + \frac{h}{2}\right) + \Phi_k^{(r)}\left(x - \frac{h}{2}\right) - 2\Phi_k^{(r)}(x) \right\}.$$

By virtue of Bernstein's inequality, we have for $|h| \leqq \delta$ the estimate

$$|\Delta^2 P| \leqq \frac{\delta^2}{4} \sum_{k=1}^{m-1} \sup_{-\infty < x < \infty} |\Phi_k^{(r+2)}(x)| \leqq \frac{\delta^2}{4} BT_0 \sum_{k=1}^{m-1} 2^k < \frac{BT_0}{4}\delta 2^m \leqq C_3\delta.$$

This proves the theorem.

A comparison of Theorem 3 with the corresponding direct theorem (Section 89) leads to the following result: *In order that the inequality*

$$\mathcal{E}_T[f] \leqq \frac{A}{T^{r+1}} \tag{3}$$

be satisfied for an arbitrary $T \geq T_0$, it is necessary and sufficient that a relation

$$f(x) = F_{T_0}(x) + \varphi(x)$$

exist, where $F_{T_0}(x) \in G_{T_0}$ and $\varphi(x)$ is a bounded function having a bounded r-th derivative belonging to Λ^*.

If we desire to replace the inequality (3) by the relation

$$\lim_{T\to\infty} T^{r+1}\mathcal{E}_T[f] = 0$$

we must change the hypothesis $\varphi^{(r)}(x) \in \Lambda^*$ to $\varphi z^{(r)}(x) \in \lambda^*$.

If the bounded function $\varphi(x)$ belongs to Λ^*, the theorems proved above give us the estimate

$$\omega(\delta;\varphi) \leqq M\delta \ln \frac{1}{\delta}.$$

92. Priwaloff's Theorem [78]. As a simple application of the results previously established we will demonstrate the following theorem proved by I. I. Priwaloff, which says: *If a periodic function $f(x)$ lies in* Lip α, *then $\tilde{f}(x)$ $(0 < \alpha < 1)$ also belongs to* Lip α. *When $\alpha = 1$, $\tilde{f}(x)$ has a modulus of continuity which does not exceed $M\delta \ln \frac{1}{\delta}$.*

We will prove at the same time a somewhat more general theorem: *Let $f(x)$ $(-\infty < x < \infty)$ be a bounded function*[10] *belonging either to* Lip α $(0 < \alpha < 1)$ *or to Λ^*. Suppose, in addition, there exists a constant C such that*

$$\int_0^x \{\tilde{f}(t) - C\}\, dt$$

is likewise bounded. Then $\tilde{f}(x)$ also is bounded and lies, in the first case in Lip α, *in the second case in Λ^*.*

Let us consider the function

$$f_1(x) = \int_0^x \{\tilde{f}(t) - C\}\, dt,$$

where C is the constant appearing in the statement of the theorem. Since $f_1(x)$ has a derivative whose conjugate belongs to Lip α or to Λ^*, as the case may be, it follows from the theorem in Section 89 that, for any positive value of T, we have the estimate

$$\mathcal{E}_T[f_1] \leqq \frac{3}{T}\,\omega\left(\frac{1}{T};f\right) \leqq \frac{A}{T^{1+\alpha}}$$

in the first case, and

$$\mathcal{E}_T[f_1] \leqq \frac{A}{T^2}$$

in the second case. By virtue of the converse theorem in Section 91, we then have

$$f_1(x) = F(x) + \varphi(x),$$

where $F(x)$ belongs to G_1 (when we put $T_0 = 1$) and $\varphi(x)$ is a bounded function having a bounded derivative which belongs to Lip α in the first case and to Λ^* in the second case.

---

[10] Consequently $\tilde{f}(x)$ exists and belongs to W^2.

Since by hypothesis $f_1(x)$ is bounded, it follows that $F(x)$ also is bounded. Consequently $F(x)$ and $F'(x)$ lie in B_1.

Accordingly,

$$\tilde{f}(x) = C + F'(x) + \varphi'(x).$$

Whence it follows that the theorem is established, because, as we already know, $F'(x)$ has a bounded derivative and hence belongs to Lip 1.

93. Generalizations of Bernstein's Theorems to the Space $L^p(p \geqq 1)$. It is to be understood that L^p denotes either the space $L^p(-\infty, \infty)$ or the space $L^p(-\pi, \pi)$ of periodic functions. *If, for an arbitrary $T \geqq T_0 > 0$, there exists in G_T a function*[11] *$F_T(x)$ with the property*

$$\|f - F_T\|_p \leqq \frac{A}{T^{r+\alpha}},$$

where A, the integer $r \geqq 0$ and $\alpha (0 < \alpha \leqq 1)$ are all three independent of T, then it follows that

$$f(x) = F_{T_0}(x) + \varphi(x).$$

Here $\varphi(x)$ belongs to L^p and has an r-th derivative which likewise belongs to L^p. In addition, for $0 < \alpha < 1$ we have $\varphi^{(r)}(x) \in \mathrm{Lip}(\alpha, p)$ and for $\alpha = 1$

$$\omega_p(\delta; \varphi^{(r)}) \leqq M\delta \ln \frac{1}{\delta}.$$

Proof: As in Section 91, let us set

$$T_k = 2^k T_0 \qquad (k = 0, 1, 2, \ldots)$$

and consider the series

$$F_{T_0}(x) + \{F_{T_1}(x) - F_{T_0}(x)\} + \{F_{T_2}(x) - F_{T_1}(x)\} + \ldots \equiv$$
$$\equiv F_{T_0}(x) + \Phi_1(x) + \Phi_2(x) + \ldots,$$

which, by virtue of our hypothesis, converges to $f(x)$ in the metric L^p.

Just as in Section 91, the inequalities

$$\|\Phi_k\|_p \leqq \frac{B}{T_k^{r+\alpha}},$$

$$\|\Phi_k^{(r)}\| \leqq \frac{B}{T_k^{\alpha}}$$

hold for $k = 1, 2, 3, \ldots$.

The second inequality shows that the series

$$\Phi_1^{r)}(x) + \Phi_2^{(r)}(x) + \ldots$$

---

[11] In the periodic case, $F_T(x)$ is a trigonometric sum.

converges in the metric L^p. In other words, there exists in L^p a function $\mathscr{K}(x)$ such that

$$\lim_{n\to\infty}\left\| \mathscr{K} - \sum_{k=1}^{n} \Phi_k^{(r)} \right\|_p = 0 . \tag{1}$$

If we now recall the relation (3) of Section 87, then we obtain the following equation:

$$\Phi_k(x) = \int_{-\infty}^{\infty} g(x-t)\ \Phi_k^{(r)}(t)\ dt + \Theta_k\ (x), \tag{2}$$

where $\Theta_k(x) \in G_c$ for a certain fixed $c < T_0$.

Since

$$\|\Theta_k\|_p \leqq \frac{B}{T_k^{r+a}} + \frac{B}{T_k^a}\int_{-\infty}^{\infty} |g(t)|\,dt$$

the series

$$\sum_{k=1}^{\infty} \|\Theta_k\|_p$$

converges. Hence there exists in L^p a function $\Theta(x)$ with the property

$$\lim_{n\to\infty}\left\| \Theta - \sum_{k=1}^{\infty} \Theta_k \right\|_p = 0. \tag{3}$$

It will be noted: With the aid of Bernstein's generalized inequality, it follows that the series

$$\sum_{k=1}^{\infty} \|\Theta_k'\|_p \leqq c \sum_{k=1}^{\infty} \|\Theta_k\|_p$$

converges. Since

$$\lim_{n\to\infty}\left\| f - F_{T_0} - \sum_{k=1}^{n} \Phi_k \right\|_p = 0$$

it follows from (1), (2), (3) that the equation

$$f(x) - F_{T_0}(x) = \int_{-\infty}^{\infty} g(t)\,\mathscr{K}(x-t)\,dt + \Theta(x)$$

is satisfied almost everywhere. Let us set

$$\varphi(x) = \psi(x) + \Theta(x)$$

with

$$\psi(x) = \int_{-\infty}^{\infty} g(t)\, \mathscr{K}(x-t)\, dt.$$

In order to complete the proof, it suffices to set up the generalized moduli of continuity of $\Theta(x)$, $\psi(x)$ respectively.

From the equation

$$\Theta_k(x+h) - \Theta_k(x) = \int_0^h \Theta_k'(x+t)\, dt$$

we get the inequality

$$\| \Theta_k(x+h) - \Theta_k(x) \|_p \leqq |h| \, \| \Theta_k' \|_p .$$

Whence we obtain the relation

$$\| \Theta(x+h) - \Theta(x) \|_p \leqq |h| \sum_{k=1}^{\infty} \| \Theta_k' \|_p = C_0 |h|,$$

that is to say, $\Theta(x)$ lies in Lip $(1, p)$.

On the other hand,

$$\omega_p(\delta; \psi) \leqq \omega_p(\delta; \mathscr{K}) \int_{-\infty}^{\infty} |g(t)|\, dt.$$

Let δ be a preassigned arbitrary positive number not exceeding 1; determine from the inequality

$$2^{m-1} \leqq \frac{1}{\delta} < 2^m$$

an integer m, as we did in Section 91; and set

$$P(x) = \sum_{k=1}^{m-1} \Phi_k^{(r)}(x),$$

$$Q(x) = \mathscr{K}(x) - P(x).$$

Then we have

$$\| Q \|_p \leqq \sum_{k=m}^{\infty} \| \Phi_k^{(r)} \|_p \leqq \frac{B}{T_0^a} \sum_{k=m}^{\infty} \frac{1}{2^{ka}} < C_1 \delta^a.$$

Consequently the inequality

$$\| Q(x+h) - Q(x) \|_p < 2C_1 \delta^a$$

holds for $|h| \leq \delta$. An estimate of the quantity

$$\| P(x+h) - P(x) \|_p ,$$

the analogue of which is contained in Section 91, gives

$$\|P(x+h)-P(x)\|_p \leqq \begin{cases} C_2\delta^\alpha & (\text{for } 0<\alpha<1), \\ \dfrac{B}{\ln 2}\,\delta \ln\dfrac{1}{\delta} & (\text{for } \alpha=1). \end{cases}$$

Accordingly, $\mathscr{K}(x) \in \mathrm{Lip}\,(\alpha, p)$ if $0<\alpha<1$, and for $\alpha=1$ we get

$$\omega_p(\delta;\mathscr{K}) \leqq M\delta \ln\frac{1}{\delta},$$

which was to be proven.

As an example, somewhat analogous to that given in Section 91, we cite here the continuous periodic function [79]

$$f(x)=\sum_{k=1}^{\infty}\frac{1}{k^{\frac{3}{2}}}e^{ikx}$$

and set $p=2$.

Since the Fourier sums furnish the best approximation in $L^2(-\pi,\pi)$, we set

$$s_{n-1}(x)=\sum_{k=1}^{n-1}\frac{1}{k^{\frac{3}{2}}}e^{ikx}.$$

Then we have

$$\|f-s_{n-1}\|_2=\sqrt{2\pi}\left\{\sum_{k=n}^{\infty}\frac{1}{k^3}\right\}^{\frac{1}{2}}<\sqrt{2\pi}\left\{\int_n^{\infty}\frac{dx}{x^3}\right\}^{\frac{1}{2}}=\frac{\sqrt{\pi}}{n}.$$

At the same time, when $0<h<\dfrac{1}{2}$, we obtain the estimate

$$\|f(x+h)-f(x-h)\|_2=2\sqrt{2\pi}\,h\left\{\sum_{k=1}^{\infty}\frac{1}{k}\left(\frac{\sin kh}{kh}\right)^2\right\}^{\frac{1}{2}}>$$

$$>3h\left\{\sum_{k<\frac{\pi}{2h}}\frac{1}{k}\right\}^{\frac{1}{2}}>3h\sqrt{\ln\frac{1}{2h}}.$$

Consequently the function under consideration does not belong to Lip (1, 2).

It will be observed that also the other results of Section 91 can be carried over into the space L^p without difficulty. So can Priwaloff's theorem. Then Λ_p (or else λ_p) would replace $L\,(1, p)$.

94. The Best Harmonic Approximation of Analytic Functions [80]. Let A denote the totality of all the analytic functions $F(w)$ $(w=t+iu)$ which

a) are real-valued on the real axis ($u = 0$),

b) are regular in the strip

(1) $$-\delta < u < \delta$$

and satisfy the inequality

$$-1 < \Re F(w) < 1.$$

If $F(w)$ belongs to the class A, then the function

$$h(t, u) = \Re F(t + iu) \quad (-\delta < u < \delta)$$

approaches a limit as $u \to \pm \delta$, for almost all values of t, and we have

$$\lim_{u \to \delta} h(t, u) = \lim_{u \to -\delta} h(t, u).$$

This limit will be denoted by $h(t)$.

It is not difficult to express the function $F(t)$ $(-\infty < t < \infty)$ as an integral over the function $h(s)$. To that end, we map the strip (1) on a half-plane by means of the transformation

$$z = ie^{\frac{\pi w}{2\delta}} \quad (z = x + iy)$$

and apply Schwarz's formula for the half-plane. Thence we obtain

(2) $$F(t) = \frac{1}{2\delta} \int_{-\infty}^{\infty} \frac{h(s)}{\operatorname{ch} \frac{\pi (t - s)}{2\delta}} ds.$$

If we allow $h(s)$ to run through the totality of all measurable functions for which the inequality

$$\operatorname*{true\,max}_{-\infty < s < \infty} |h(s)| \leq 1$$

is fulfilled, we obtain the class A_0 of all those real-valued functions of a real argument $t(-\infty < t < \infty)$ which are identical with any function within A, restricted to the real axis.

We are now going to determine the quantity $\mathcal{E}_T\{A_0\}$ (cf. Section 85).

The function

$$g(t) = \frac{1}{2\delta} \frac{1}{\operatorname{ch} \frac{\pi t}{2\delta}}$$

appears as the kernel in the integral relation (2). Now, since the function

$$\lambda(t) = \frac{1}{2\delta} \int_{-\infty}^{\infty} \frac{\cos tv}{\operatorname{ch} \frac{\pi v}{2\delta}} dv = \frac{1}{\operatorname{ch} \delta t}$$

is triply monotonic when t is sufficiently large, we can deduce by the application of Nagy's criterion (Section 86) that $g(t)$ is a Krein kernel.

Hence, for sufficiently large arbitrary values of T, we have the equation

$$\mathcal{E}_T\{A_0\} = \frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k \frac{1}{2k+1}\frac{1}{\operatorname{ch}(2k+1)T\bar{\delta}}.$$

If $F(t)$ belongs to A_0 and has the period 2π, then it follows from the foregoing that, for sufficiently large values of n, the deviation from $F(t)$ of the best approximation through trigonometric polynomials of order less than the n-th, in the metric of C, is equal to

$$\varrho_n \leqq \frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k \frac{1}{2k+1}\frac{1}{\operatorname{ch}(2k+1)n\delta}.$$

This result can also be easily established at once. From the foregoing proof it follows, among other things, that the above inequality is valid for all values of n.

To prove the last inequality, let us map the strip (1) of the w-plane on the circular ring

$$q = e^{-\delta} < |z| < e^{\delta} = \frac{1}{q}$$

of the z-plane, by means of the transformation

$$z = e^{iw}.$$

The function

$$F\left(\frac{1}{i}\ln z\right)$$

is regular in this circular ring and is evidently also single-valued, since we have assumed here that $F(w)$ is periodic.

With the help of the Laurent development

$$\sum_{m=-\infty}^{\infty} c_m z^m$$

we find without difficulty that

$$c_m = \bar{c}_{-m} = \frac{1}{\pi}\frac{q^m}{1+q^{2m}}\int_0^{2\pi} e^{-im\Theta}h(\Theta)\,d\Theta.$$

Hence it follows that

$$F(t) = \frac{1}{2\pi}\int_0^{2\pi}\left\{1 + 4\sum_{m=1}^{\infty}\frac{q^m}{1+q^{2m}}\cos m(t-\Theta)\right\}h(\Theta)\,d\Theta. \tag{3}$$

It will be observed that the representation (3) can be obtained immediately from (2) if we take into consideration the periodicity of $h(s)$.

In the theory of elliptic functions we find the following formulas:

$$1 + 4\sum_{m=1}^{\infty} \frac{q^m}{1+q^{2m}} \cos mv = \frac{2\mathbf{K}}{\pi} \operatorname{dn} \frac{\mathbf{K}v}{\pi}, \tag{4}$$

$$\operatorname{dn} \frac{\mathbf{K}v}{\pi} = \sqrt{\varkappa'} \prod_{m=1}^{\infty} \frac{1 + 2q^{2m-1}\cos v + q^{4m-2}}{1 - 2q^{2m-1}\cos v + q^{4m-2}}. \tag{5}$$

By reason of the formula (4), the representation (3) can be written in the form

$$F(t) = \frac{\mathbf{K}}{\pi^2} \int_0^{2\pi} \operatorname{dn} \frac{\mathbf{K}v}{\pi} h(t+v)\, dv .$$

Whence it follows that, no matter how the constants ζ_k are chosen, we always have

$$\frac{\mathbf{K}}{\pi^2} \int_0^{2\pi} h(t+v) \left\{ \operatorname{dn} \frac{\mathbf{K}v}{\pi} - \sum_{m=0}^{n-1} \zeta_m \cos mv \right\} dv =$$

$$= F(t) - \frac{\mathbf{K}}{\pi} \left\{ 2\zeta_0 a_0 + \sum_{m=1}^{n-1} \zeta_m (\alpha_m \cos mt + \beta_m \sin mt) \right\}$$

where α_m, β_m are the Fourier coefficients of $h(\Theta)$.

Since

$$\sum_{m=0}^{n-1} \zeta_m \cos mv = \sum_{m=0}^{n-1} z_m \cos^{2m} \frac{v}{2}$$

and since the inequality

$$|h(\Theta)| \leqq 1$$

is satisfied almost everywhere, it follows that

$$\max_{0 \leqq t \leqq 2\pi} \left| F(t) - \frac{\mathbf{K}}{\pi} \left\{ 2\zeta_0 a_0 + \sum_{m=1}^{n-1} \zeta_m (\alpha_m \cos mt + \beta_m \sin mt) \right\} \right| \leqq$$

$$\leqq \frac{2\mathbf{K}}{\pi^2} \int_0^{\pi} \left| \operatorname{dn} \frac{\mathbf{K}v}{\pi} - \sum_{m=0}^{n-1} z_m \cos^{2m} \frac{v}{2} \right| dv$$

and accordingly

$$\varrho_n \leqq \min_{z_m} \frac{2\mathbf{K}}{\pi^2} \int_0^{\pi} \left| \operatorname{dn} \frac{\mathbf{K}v}{\pi} - \sum_{m=0}^{n-1} z_m \cos^{2m} \frac{v}{2} \right| dv .$$

By reason of the considerations in Section 51, we can assert that the right-hand side is equal to

$$\frac{2\mathbf{K}}{\pi^2}\int_0^{\pi} \operatorname{dn}\frac{\mathbf{K}v}{\pi}\cdot \operatorname{sign}\cos nv\,dv = \frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k\frac{1}{2k+1}\frac{1}{\operatorname{ch}(2k+1)n\delta}$$

provided we succeed in proving that, for appropriately chosen coefficients z_m, the difference

$$\operatorname{dn}\frac{\mathbf{K}v}{\pi} - \sum_{m=0}^{n-1} z_m \cos^{2m}\frac{v}{2} \tag{6}$$

changes sign only at the points where the function $\cos nv$ does. We will prove that the difference (6) cannot change sign more than n times in the interval $0 < v < \pi$ for any choice of the coefficients.

In fact, by virtue of (5) we have

$$\operatorname{dn}\frac{\mathbf{K}v}{\pi} = \sqrt{\varkappa'}\prod_{m=0}^{\infty}\left\{\frac{\operatorname{ch}(2m-1)\delta}{\operatorname{ch}^2\left(m-\frac{1}{2}\right)\delta - \cos^2\frac{v}{2}} - 1\right\}.$$

Therefore all the derivatives of $\operatorname{dn}\dfrac{\mathbf{K}v}{\pi}$ with respect to the variable $\cos^2\dfrac{v}{2}$ are positive. Consequently the n-th derivative of the function (6) with respect to $\cos^2\dfrac{v}{2}$ is positive, and our assertion is proved.

That the equality sign in the estimate of ϱ_n is acceptable, is shown by the "extremal" function $F_0(w)$, for which

$$h_0(t) = \operatorname{sign}\cos nt\,.$$

It will be observed, in addition, that from the results derived by us we have

$$\varrho_n < \frac{8}{\pi}q^n\,. \tag{7}$$

95. A Different Formulation of the Result of the Preceding Section. Let $f(z)$ $(z = x + iy)$ be an analytic function which

α) is regular in the interior of an ellipse having the foci ± 1 and the sum of the semi-axes equal to $\dfrac{1}{q}$;

β) is real-valued along the major axis of this ellipse;

γ) satisfies the inequality

$$-1 < \Re f(z) < 1$$

in the interior of this ellipse.

By means of the transformation

$$z = \cos w$$

we obtain an even function of w

$$F(w) = f(z);$$

which belongs to the class A of the previous section and has the period 2π, whereby

$$\delta = \ln \frac{1}{q} .$$

To the segment $[-1, 1]$ in the z-plane, there now corresponds the real axis in the w-plane. Since the aforestated transformation changes a polynomial of the $(n-1)$st degree in x into an even trigonometric sum of the $(n-1)$st order in t, we obtain the following theorem: *If a function $f(z)$ satisfies the conditions $\alpha)$, $\beta)$, $\gamma)$, then the deviation $\mathsf{E}_{n-1}[f]$ of the best approximation, from $f(x)$, through trigonometric polynomials of the $(n-1)$st degree, in the interval $[-1, 1]$, satisfies the inequality*

$$\mathsf{E}_{n-1}[f] \leqq \frac{8}{\pi} \sum_{m=0}^{\infty} (-1)^m \frac{1}{2m+1} \frac{q^{(2m+1)n}}{1+q^{2(2m+1)n}} .$$

This inequality cannot be sharpened if nothing more is known about $f(z)$.

As formulated here, the above result is a sharpening of an important theorem due to Bernstein [67], to the effect that

$$\lim_{n\to\infty} \sqrt[n]{\mathsf{E}_n[f]} \leqq q.$$

It is easy to see that this inequality still remains true if the conditions $\beta)$, $\gamma)$ are no longer fulfilled.

In fact: Let us take a smaller confocal ellipse the sum of whose semiaxes is $\frac{1}{r}$ $(r > q)$ and then split up $f(z)$ so that

$$f(z) = f_1(z) + i f_2(z),$$

where $f_1(z)$, $f_2(z)$ are two functions satisfying the conditions $\alpha)$, $\beta)$. Thus we obtain by virtue of the formula (7), Section 94,

$$\mathsf{E}_{n-1}[f] \leqq \mathsf{E}_{n-1}[f_1] + \mathsf{E}_{n-1}[f_2] < \frac{8}{\pi} r^n \{M_1(r) + M_2(r)\},$$

where $M_k(r)$ is the greatest absolute value of $\Re f_k(z)$ $(k = 1, 2)$ on the smaller ellipse. Therefore

$$\overline{\lim_{n\to\infty}} \sqrt[n]{\mathsf{E}_n[f]} \leqq r,$$

and, since r can be chosen as near to q as we please, it follows that

$$\lim_{n\to\infty} \sqrt[n]{\mathsf{E}[f]} \leqq q.$$

Similarly, the following theorem can be derived from the results of Section 94: *If $F(t)$ can be brought into the form*

$$F(t) = \Phi(t) + \varphi(t)$$

where $\Phi(t) \in G_T$ and $\varphi(w)$ $(w = t + iu)$ is regular in the strip

$$-\delta < u < \delta$$

and bounded in every smaller strip

$$-\delta' \leqq u \leqq \delta' \qquad (0 < \delta' < \delta),$$

then

$$\lim_{T \to \infty} \{\mathcal{E}_T[F]\}^{\frac{1}{T}} \leqq e^{-\delta}.$$

In order to elucidate this theorem, let us recall Sections 37 and 38. There it was proved that for $a > 1$ the relations

$$\mathsf{E}_n\left[\frac{1}{x-a}\right] = \frac{(a-\sqrt{a^2-1})^n}{a^2-1}, \tag{1}$$

$$\mathsf{E}_n\left[\frac{A(a^2-1)^2}{(x-a)^2} + \frac{A'(a^2-1)}{x-a}\right] = \frac{1}{2}|A|\frac{\sqrt{a^2-1}}{(a+\sqrt{a^2-1})^n}\left|n + \frac{2aA-A'}{A\sqrt{a^2-1}}\right| + \tag{2}$$

$$+ \sqrt{\left(n + \frac{2aA-A'}{A\sqrt{a^2-1}}\right)^2 + \frac{1}{a^2-1}}\,(1+\varepsilon_n)$$

are valid, where $\lim_{n\to\infty} \varepsilon_n = 0$. Each of the functions

$$\frac{1}{x-a}, \qquad \frac{A(a^2-1)^2}{(x-a)^2} + \frac{A'(a^2-1)}{x-a}$$

satisfies the conditions $\alpha)$, $\beta)$ if we put

$$q = a - \sqrt{a^2-1} = \frac{1}{a+\sqrt{a^2-1}}.$$

In fact: for the ellipse of regularity

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

with the foci ± 1, the sum of the semi-axes is equal to

$$a + b = a + \sqrt{a^2-1}.$$

On the basis of the estimates derived in this section, we can conclude merely that

$$\overline{\lim_{n\to\infty}} \sqrt[n]{E_n\left[\frac{1}{x-n}\right]} \leqq a - \sqrt{a^2-1},$$

$$\overline{\lim_{n\to\infty}} \sqrt[n]{E_n\left[\frac{A(a^2-1)^2}{(x-a)^2} + \frac{A'(a^2-1)}{x-a}\right]} \leqq a - \sqrt{a^2-1}\ .$$

Thus we see that the relations (1), (2) in Sections 37, 38 are incomparably sharper.

96. The Converse of Bernstein's Theorem [67], [69] says: *If $F(t)$ is a function which is defined along the entire number axis and such that*

$$\overline{\lim_{T\to\infty}} \{\mathcal{E}_T[F]\}^{\frac{1}{T}} \leqq e^{-\delta} \quad (\delta > 0)\ , \tag{1}$$

then $F(t)$ can be brought into the form

$$F(t) = \Phi(t) + \varphi(t)\ ,$$

where $\Phi(t) \in G_\tau$ and $\varphi(w)$ $(w = t + iu)$ is regular in the strip

$$-\delta < u < \delta$$

and bounded in every smaller strip

$$-\delta' \leqq u \leqq \delta' \qquad (0 < \delta' < \delta).$$

This theorem, combined with the theorems of the preceding Section, enables us to determine the size of the strip in which $F(w)$ is regular, by means of $\mathcal{E}_T[F]$. It is reminiscent of the Cauchy-Hadamard theorem, which allows us to determine the radius of convergence of a power series from the coefficients. Half the width of the strip of regularity is in fact equal to

$$\ln \frac{1}{\overline{\lim_{T\to\infty}} \{\mathcal{E}_T[F]\}^{\frac{1}{T}}}.$$

Similarly, half the sum of the axes of the ellipse of regularity is given by

$$\frac{1}{\overline{\lim_{n\to\infty}} \sqrt[n]{E_n[f]}}\ .$$

To prove this theorem let us take any positive $\delta_1 < \delta$ and determine, in conformity with the condition (1), a number τ with the property: For an arbitrary $T \geqq \tau$ there exists in G_T a function $\Phi_T(t)$ which satisfies the inequality

$$\sup_{-\infty<t<\infty} |F(t) - \Phi_T(t)| \leqq e^{-T\delta_1} \tag{2}$$

We then set

$$\Phi(t) = \Phi_\tau(t), \quad \varphi_k(t) = \Phi_{\tau+k}(t) - \Phi_{\tau+k-1}(t) \qquad (k = 1, 2, \ldots)$$

and obtain by virtue of (2) the uniformly convergent development

$$F(t) = \Phi(t) + \sum_{k=1}^{\infty} \varphi_k(t).$$

Thus we obtain from (2):

$$\sup_{-\infty < t < \infty} |\varphi_k(t)| \leqq 2e^{-(\tau+k-1)\delta_1},$$

that is to say $\varphi_k(w)$ is bounded along the real axis and hence belongs to $B_{\tau+k}$.

Whence it follows, as we have shown in Section 75, that for any $w = t + iu$ the inequality

$$|\varphi_k(w)| \leqq 2e^{-(\tau+k-1)\delta_1} e^{(\tau+k)|u|}$$

is valid. Therefore the series

$$\sum_{k=1}^{\infty} \varphi_k(w) \tag{3}$$

is absolutely and uniformly convergent in the half-strip

$$-\delta' \leqq u \leqq \delta' \tag{4}$$

with any positive $\delta' < \delta_1$. Hence the sum of the series (3) is a function which is regular in the strip

$$-\delta_1 < u < \delta_1 \tag{5}$$

and will be denoted by $\varphi(w)$. This function is bounded on every part of the strip (4):

$$|\varphi(w)| \leqq 2e^{\delta_1} \sum_{k=1}^{\infty} e^{-(\tau+k)(\delta-\delta_1')} \leqq \frac{2e^{\delta_1}}{e^{\delta_1-\delta'}-1}.$$

We will now show that $\varphi(w)$ is regular, not only in the half-strip (5) but also in the strip

$$-\delta < u < \delta \tag{6}$$

and is bounded in the half-strip (4), not only for $\delta' < \delta_1$ but also for any $\delta' < \delta$. For that purpose let us take any δ_1^* satisfying the inequality

$$\delta_1 < \delta_1^* < \delta$$

and repeat the above method of reasoning, now for δ_1^*. Then we obtain a representation of the function $F(t)$

$$F(t) = \Phi^*(t) + \varphi^*(t),$$

where $\Phi^*(t) \in G_{\tau^*}$ and $\varphi^*(w)$ is regular in the half-strip

$$-\delta_1^* < u < \delta_1^* \tag{7}$$

and bounded in the half-strip (4) for any $\delta' < \delta_1^*$. Now, since the relation

$$\Phi^*(t) + \varphi^*(t) = \Phi(t) + \varphi(t)$$

is valid, the function

$$\varphi(w) = \varphi^*(w) + \{\Phi^*(w) - \Phi(w)\}$$

also is regular in the half-strip (7) as well as in (6), inasmuch as δ_1^* can be chosen as near to δ as we please.

Furthermore, since the difference $\varphi^*(t) - \varphi(t)$ is bounded along the entire real axis, it follows that $\Phi^*(w) - \Phi(w)$ belongs to the class B_σ if we set $\sigma = \max\{\tau, \tau^*\}$. Therefore the difference $\Phi^*(w) - \Phi(w)$ is bounded in every strip parallel to the real axis. Consequently the function $\varphi(w)$ is bounded in the half- strip (4) for any $\delta' < \delta_1^*$ and hence for $\delta' < \delta$.

This proves the theorem.

CHAPTER VI

WIENER'S THEOREM ON APPROXIMATION

97. Wiener's Problem [81]. Given a certain set $\mathfrak{M}$ in the space $L(-\infty, \infty)$. Let us consider all possible functions of the form

$$\sum_{\alpha,\beta} c_{\alpha,\beta} f_\alpha(x + \lambda_{\alpha,\beta}), \tag{1}$$

where $c_{\alpha,\beta}$ are any numbers, $\lambda_{\alpha,\beta}$ are any real numbers and f_α lies in $\mathfrak{M}$. Every function of the form (1) lies in L and the totality of these functions constitutes a certain linear manifold in the space L. The closure of this set will be denoted by $I_{\mathfrak{M}}$.

Wiener's problem is: To find necessary and sufficient conditions for the set $\mathfrak{M}$ so that $I_{\mathfrak{M}} = L$. In other words: It is required to find conditions under which, for an arbitrary function $g(x) \in L$ and an arbitrary $\varepsilon > 0$, there exists a function $h(x)$ of the form (1) such that:

$$\int_{-\infty}^{\infty} |g(x) - h(x)|\, dx < \varepsilon.$$

Wiener's result can be formulated simply: $I_{\mathfrak{M}} = L$ *if and only if there does not exist any point* x_0 $(-\infty < x_0 < \infty)$ *at which the Fourier transforms of all functions in* $\mathfrak{M}$ *become zero.*

The proof of this important theorem will engage our attention in several sections.

98. The Necessity of Wiener's Condition is proved very simply. For, suppose a point x_0 $(-\infty < x_0 < \infty)$ exists such that

$$\int_{-\infty}^{\infty} f(t)\, e^{-itx_0}\, dt = 0$$

for an arbitrary function $f(t) \in \mathfrak{M}$.

Let us choose a function $g(t) \in L$ for which

$$\int_{-\infty}^{\infty} g(t)\, e^{-itx_0}\, dt = 1\ .$$

Then, for arbitrary $c_{\alpha,\beta}$ and real $\lambda_{\alpha,\beta}$, we have

$$\int_{-\infty}^{\infty} \left| g(t) - \sum_{\alpha,\beta} c_{\alpha,\beta} f_\alpha(t + \lambda_{\alpha,\beta}) \right| e^{-ix_0 t}\, dt = 1 - \sum_{\alpha,\beta} c_{\alpha,\beta}\, e^{i\lambda_{\alpha,\beta} x_0} \int_{-\infty}^{\infty} f_\alpha(u)\, e^{-iux_0}\, du = 1$$

and

$$\int_{-\infty}^{\infty} \left| g(t) - \sum_{\alpha,\beta} c_{\alpha,\beta} f_\alpha(t + \lambda_{\alpha,\beta}) \right| dt \geqq 1,$$

that is to say $g(t)$ does not belong to $I_{\mathfrak{M}}$.

The proof of the sufficiency of Wiener's condition is rather complicated and requires a number of preliminary considerations.

99. Some Definitions and Notation. We agree to let $U_\lambda (-\infty < \lambda < \infty)$ denote the operator which carries $f(x) \in L$ into $f(x - \lambda)$. This operator does not alter the element norm

$$\| U_\lambda f \| = \int_{-\infty}^{\infty} | f(x + \lambda) |\, dx = \int_{-\infty}^{\infty} | f(x) |\, dx = \| f \| .$$

Furthermore, from the considerations in Section 80,

$$\lim_{\lambda \to 0} \| U_\lambda f - f \| = \lim_{\lambda \to 0} \int_{-\infty}^{\infty} | f(x + \lambda) - f(x) |\, dx = 0. \tag{1}$$

We know that the Faltung

$$f * g = \int_{-\infty}^{\infty} f(u)\, g(t - u)\, du$$

of two functions $f, g \in L$ lies also in L; therefore

$$\| f * g \| \leqq \| f \| \, \| g \| . \tag{2}$$

We agree further that the expression

$$f^* = f^*(t) = \int_{-\infty}^{\infty} f(u)\, e^{-iut}\, du \quad (f \in L)$$

shall denote the Fourier transform without the factor $\frac{1}{\sqrt{2\pi}}$. Further, $\mathfrak{N}^*$ will be understood to mean the set through which f^* runs if f runs through the set $\mathfrak{N} \subset L$. In particular, $\mathfrak{N}$ can coincide with L and, in that case, we obtain a linear

space $L^* \subset C_\infty$. Since two distinct functions f, $g \in L$ correspond to distinct functions f^*, g^* because of the theorem in Section 65, we can introduce a norm into L by setting

$$\|f^*\| = \int_{-\infty}^{\infty} |f(t)|\, dt = \|f\| \ .$$

It is easy to see that the space L^* so normalized is complete.

Definition 1. A complete linear normalized space K is called a *normalized ring* K if a multiplication of the elements in K is so defined that the following requirements are fulfilled:

1) $f \cdot (\lambda g + \mu h) = \lambda f \cdot g + \mu f \cdot h$,
$(\lambda g + \mu h) \cdot f = \lambda g \cdot f + \mu h \cdot f$ $\qquad (f, g, h \in K;\quad \lambda, \mu$ are numbers)

2) $(f \cdot g) \cdot h = f \cdot (g \cdot h)$;

3) $\|f \cdot g\| \leqq \|f\|\, \|g\|$.

The space L is a normalized ring if the product of two elements of the ring is defined as the Faltung of these elements. Actually, the property (3) implies the relation (2), and the remaining properties are self-evident. Hereafter the space L will often be called the ring L.

The space L^* also is a normalized ring if the product is understood to be the ordinary algebraic product. In fact, if f^*, $g^* \in L^*$, then (cf. Section 66)

$$f^* g^* = \{f * g\}^*$$

and consequently

$$\|f^* g^*\|_{L^*} = \|f * g\|_L \leqq \|f\|_L\, \|g\|_L = \|f^*\|_{L^*}\, \|g^*\|_{L^*} .$$

Definition 2. A specified set of elements of the normalized ring K is called an *ideal* in K if

1. the set contains, together with the elements f and g, also

$$p \cdot f + q \cdot g$$

for any $p, q \in K$;

2. $f_n \in I$ and $\|f - f_n\| \to 0$ imply: $f \in I$.

We will prove that the set $I_{\mathfrak{M}}$, introduced in Section 97, is an ideal [1] in the ring L.

---

[1] This justifies our designation.

Let $f \in I_{\mathfrak{M}}$, $g \in L$. Then we have

$$\left\| \int_{-\infty}^{\infty} f(t-u)\, g(u)\, du - \sum_{k=-\infty}^{\infty} f(t-k\lambda) \int_{k\lambda}^{(k+1)\lambda} g(u)\, du \right\| =$$

$$= \left\| \sum_{k=-\infty}^{\infty} \int_{k\lambda}^{(k+1)\lambda} \{f(t-u) - f(t-k\lambda)\}\, g(u)\, du \right\| \leqq$$

$$\leqq \sum_{k=-\infty}^{\infty} \int_{k\lambda}^{(k+1)\lambda} |g(u)|\, \|U_{u-k\lambda} f - f\|\, du \leqq \|g\| \max_{0 \leqq \delta \leqq \lambda} \|U_\delta f - f\|.$$

Since

$$\max_{0 \leqq \delta \leqq \lambda} \|U_\delta f - f\| = \max_{0 \leqq \delta \leqq \lambda} \int_{-\infty}^{\infty} |f(t+\delta) - f(t)|\, dt = \omega_1(\lambda)$$

approaches zero as $\lambda \to 0$ (see Section 80), it follows that, for sufficiently small values of $\lambda = \lambda(\varepsilon)$

$$\left\| \int_{-\infty}^{\infty} f(t-u)\, g(u)\, du - \sum_{k=-\infty}^{\infty} f(t-k\lambda) \int_{k\lambda}^{(k+1)\lambda} g(u)\, du \right\| \leqq \frac{\varepsilon}{2}.$$

On the other hand, when λ is fixed and N is large enough, we have

$$\left\| \sum_{k \geqq N+1} f(t-k\lambda) \int_{k\lambda}^{(k+1)\lambda} g(u)\, du \right\| \leqq \|f\| \left\{ \int_{N\lambda}^{\infty} |g(u)|\, du + \int_{-\infty}^{-N\lambda} |g(u)|\, du \right\} \leqq \frac{\varepsilon}{2}.$$

Therefore

$$\left\| f * g - \sum_{k=-N}^{N} c_k f(t-k\lambda) \right\| \leqq \varepsilon$$

with

$$c_k = \int_{k\lambda}^{(k+1)\lambda} g(u)\, du.$$

Since, moreover, $I_{\mathfrak{M}}$ is a closed set, it follows that $f * g \in I_{\mathfrak{M}}$. But that is what we were to prove.

It is easy to see that $I^*_{\mathfrak{M}}$ is an ideal in the ring L^*.

100. Several Lemmas. Let us consider the function

$$\mu(t) = \frac{\cos t - \cos 2t}{\pi t^2},$$

which lies in L. We have already seen (cf. Section 90):

$$\mu^*(t) = \begin{cases} 1, & \text{for } -1 \leqq t \leqq 1, \\ 2 - |t|, & \text{for } 1 \leqq |t| \leqq 2, \\ 0, & \text{for } |t| \geqq 2. \end{cases}$$

Lemma 1. If $f^*(t) \in L^*$, then

$$\lim_{a\to\infty} \left\| f^*(t) - f^*(t)\, \mu^*\left(\frac{t}{a}\right) \right\| = 0 \tag{1}$$

and

$$\lim_{\varepsilon\to 0} \left\| \{f^*(t) - f^*(u)\}\, \mu^*\left(\frac{t-u}{\varepsilon}\right) \right\| = 0 \quad (\varepsilon > 0). \tag{2}$$

Proof: By observing that

$$g^*\left(\frac{t-\beta}{a}\right) = \{a e^{i\beta t} g(at)\}^*$$

we obtain

$$f^*(t) - f^*(t)\, \mu^*\left(\frac{t}{a}\right) = f^*(t)\, \mu^*(0) - f^*(t)\, \mu^*\left(\frac{t}{a}\right) =$$

$$= \int_{-\infty}^{\infty} f(u)\, e^{-iut}\, du \int_{-\infty}^{\infty} a\mu(av)\, dv - \int_{-\infty}^{\infty} e^{-iut}\, du \int_{-\infty}^{\infty} f(u-v)\, a\mu(av)\, dv.$$

Therefore we have the relations

$$\left\| f^*(t) - f^*(t)\, \mu^*\left(\frac{t}{a}\right) \right\| = \left\| \int_{-\infty}^{\infty} a\mu(av)\, \{f(t) - f(t-v)\}\, dv \right\| =$$

$$= \left\| \int_{-\infty}^{\infty} \mu(s) \left\{ f(t) - f\left(t - \frac{s}{a}\right) \right\} ds \right\| \leqq \int_{-\infty}^{\infty} |\mu(s)|\, \| U_{\frac{s}{a}} f - f \|\, ds =$$

$$= \int_{-N}^{N} |\mu(s)|\, \| U_{\frac{s}{a}} f - f \|\, ds + \int_{-\infty}^{-N} + \int_{N}^{\infty} = I_1 + I_2 + I_3.$$

Since

$$\| U_{\frac{s}{a}} f - f \| \leqq 2 \| f \|$$

and

$$\lim_{a\to\infty} \| U_{\frac{s}{a}} f - f \| = 0 \qquad (-N \leqq s \leqq N)$$

it follows that

$$\lim_{a\to\infty} I_1 = 0$$

for every fixed N. On the other hand,

$$|I_2 + I_3| \leqq 2\,||f||\left\{\int_{-\infty}^{-N} |\mu(s)|\,ds + \int_{N}^{\infty} |\mu(s)|\,ds\right\}$$

and therefore

$$\lim_{N\to\infty}\{I_2 + I_3\} = 0.$$

Thus (1) is proved.

We have further

$$\{f^*(t) - f^*(u)\}\,\mu^*\left(\frac{t-u}{\varepsilon}\right) = \int_{-\infty}^{\infty} e^{-itx}\,dx \int_{-\infty}^{\infty} \varepsilon\mu(\varepsilon x - \varepsilon y)\,e^{-iu(x-y)} f(y)\,dy -$$

$$-\int_{-\infty}^{\infty} e^{-itx}\,\varepsilon\mu(\varepsilon x)\,e^{iux}\,dx \int_{-\infty}^{\infty} e^{-iuy} f(y)\,dy.$$

Hence

$$\left\|\{f^*(t) - f^*(u)\}\,\mu^*\left(\frac{t-u}{\varepsilon}\right)\right\| = \left\|\int_{-\infty}^{\infty} \varepsilon\,\{\mu(\varepsilon t - \varepsilon y) - \mu(\varepsilon t)\}\,e^{-iu(t-y)} f(y)\,dy\right\| \leqq$$

$$\leqq \int_{-\infty}^{\infty} \varepsilon\,||\mu(\varepsilon t - \varepsilon y) - \mu(\varepsilon t)||\,|f(y)|\,dy = \int_{-\infty}^{\infty} ||\mu(t - \varepsilon y) - \mu(t)||\,|f(y)|\,dy,$$

since, for an arbitrary $\varepsilon > 0$ and $g \in L$,

$$\varepsilon\,||g(\varepsilon t)|| = ||g(t)||.$$

If we split up the integral on the right, we find, as before, that the right-hand side approaches zero as $\varepsilon \to 0$. Thus (2) is also proved.

Lemma 2. *Let I^* be an ideal in L^*. If, to every point of the real axis, inclusive of the point at infinity, there corresponds a neighborhood* [2] Γ_u *and a function $\omega_u(t) \in I^*$ with $h(t) = \omega_u(t)$ $(t \in \Gamma_u)$, then $h(t) \in I^*$.*

Proof: According to the Heine-Borel theorem there exists a finite number of intervals Γ_n covering the number axis. Let these intervals be $\Gamma_{u_k} = [a_k, b_k]$, where Γ_{u_1} covers the point at infinity. We may assume that

$$a_k < b_{k-1} < a_{k+1} < b_k \qquad (k = 2, 3, \ldots, n;\ a_{n+1} = a_1).$$

Let us construct a piecewise continuous linear function $\lambda_k(t)$ $(k = 2, 3, \ldots, n)$, equal to zero outside the interval (a_k, b_k). Then we obtain a curve, represented in

[2] By a neighborhood of the point at infinity is meant the set of points which lie outside of a finite interval. Thus the number axis appears as the topological image of a circle; that is, among others, the reason for the applicability of the Heine-Borel theorem, which we shall utilize later.

Fig. 3, and an additional piecewise linear function $\lambda_1(t)$ which is equal to zero in the interval (b_1, a_{n+1}) and is represented in Fig. 4.

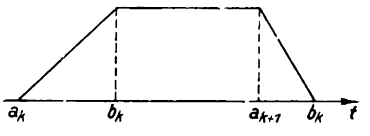


Fig. 3

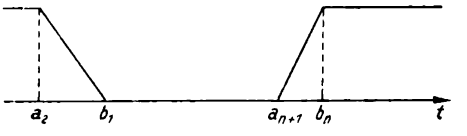


Fig. 4

It is clear that

$$\sum_{k=1}^{n} \lambda_k(t) = 1 .$$

Therefore

$$h(t) = \sum_{k=1}^{n} h(t)\, \lambda_k(t) .$$

But since $h(t) = \omega_{u_k}(t)$ when $t \in \Gamma_{u_k}$ (i.e., when $\lambda_k(t) \neq 0$), we have

$$h(t) = \sum_{k=1}^{n} \omega_{u_k}(t)\, \lambda_k(t) = \sum_{k=2}^{n} \omega_{u_k}(t)\, \lambda_k(t) + \omega_{u_1}(t)\, \{\lambda_1(t) - 1\} + \omega_{u_1}(t) .$$

However,

$$\omega_{u_k}(t) \in I^* \qquad (k = 1, 2, \ldots, n)$$

and, on the other hand,

$$\lambda_k(t) \in L^* \qquad (k = 2, 3, \ldots, n),$$
$$\lambda_1(t) - 1 \in L^*,$$

because, when $k = 2, 3, \ldots, n$, we have

$$\lambda_k(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itu} \left\{ \frac{e^{ib_{k-1}u} - e^{ia_k u}}{b_{k-1} - a_k} - \frac{e^{ib_k u} - e^{a_{k+1}u}}{b_k - a_{k+1}} \right\} \frac{du}{u^2}$$

and $\lambda_1(t) - 1$ admits an analogous representation. Consequently:

$$h(t) \in I^* .$$

101. The Wiener-Levy Theorem [82]: *If $f^*(t) \in L^*$ and $\Phi(z)$ are analytic functions which are regular on the curve $z = f^*(t)$, where $\Phi(0) = 0$, then $\Phi[f^*(t)] \in L^*$.*

Proof: Let u be any finite point on the real axis. Put $z_0 = f^*(u)$ and let ϱ denote the radius of convergence of the series

$$\Phi(z) = a_0 + \sum_{k=1}^{\infty} a_k (z - z_0)^k ,$$

which represents $\Phi(z)$ in a neighborhood of z_0.

According to statement (2) in Lemma 1 of the previous section there exists an $\varepsilon > 0$ such that

$$\delta = \left\| \{f^*(t) - f^*(u)\}\, \mu^*\left(\frac{t-u}{\varepsilon}\right) \right\| < \varrho \ ;$$

let us set

$$\gamma(t) = f^*(u) + \{f^*(t) - f^*(u)\}\, \mu^*\left(\frac{t-u}{\varepsilon}\right).$$

Then the function

$$\Phi[\gamma(t)] - a_0 = \sum_{k=1}^{\infty} a_k [\gamma(t) - z_0]^k \tag{1}$$

lies in L^*. In particular, the function

$$\gamma(t) - z_0 = \{f^*(t) - f^*(u)\}\, \mu^*\left(\frac{t-u}{\varepsilon}\right)$$

lies in L^*. Therefore L^* contains every term of the right-hand side of (1), and we have only to remember that

$$\left\| \sum_{k=1}^{\infty} a_k [\gamma(t) - z_0]^k \right\| \leqq \sum_{k=1}^{\infty} |a_k|\, \delta^k < \infty \ .$$

Together with the function

$$\Phi[\gamma(t)] - a_0$$

L^* contains the function

$$\mu^*\left(\frac{t-u}{\varepsilon}\right) \{\Phi[\gamma(t)] - a_0\}$$

and therewith also the function

$$\mu^*\left(\frac{t-u}{\varepsilon}\right) \Phi[\gamma(t)] = \omega_u(t),$$

which is equal to $\Phi[f^*(t)]$ when $u - \varepsilon \leq t \leq u + \varepsilon$.

In order to be in a position to apply Lemma 2 (for the case $I^* = L^*$), it remains merely to prove the existence of such a function $\omega_\infty(t) \in L^*$ that for a certain $a > 0$ for $|t| > 2a$ we have

$$\omega_\infty(t) = \Phi[f^*(t)] \tag{2}$$

To that end, let us keep in mind that $f^*(\pm\infty) = 0$ and consider the following development of the function $\Phi(z)$ in the neighborhood of the point $z = 0$:

$$\Phi(z) = \sum_{k=1}^{\infty} a_k z^k.$$

Let us take any finite point of the real axis. By hypothesis, there exists a function $f(t) \in \mathfrak{M}$ such that

$$f^*(u) \neq 0 \ .$$

Now consider the identity

$$f^*(u)\, \mu^*\left(\frac{t-u}{\varepsilon}\right) = \mu^*\left(\frac{t-u}{\varepsilon}\right) f^*(t) - \mu^*\left(\frac{t-u}{\varepsilon}\right)\{f^*(t) - f^*(u)\}.$$

The first term on the right-hand side surely lies in $I^*_{\mathfrak{M}}$. Hence we must merely prove that this property belongs also to the second term. Now, for every value of t

$$\mu^*\left(\frac{t-u}{\varepsilon}\right)\{f^*(t) - f^*(u)\} =$$

$$= \mu^*\left(\frac{t-u}{\varepsilon}\right) \frac{\mu^*\left(\frac{t-u}{2\varepsilon}\right)\{f^*(t) - f^*(u)\}}{\mu^*\left(\frac{t-u}{2\varepsilon}\right)\{f^*(t) - f^*(u)\} + f^*(u)} f^*(t) =$$

$$= \mu^*\left(\frac{t-u}{\varepsilon}\right) f^*(t)\, \frac{h^*(t)}{1 + h^*(t)} \ ,$$

where

$$h^*(t) = \frac{\mu^*\left(\frac{t-u}{2\varepsilon}\right)\{f^*(t) - f^*(u)\}}{f^*(u)} \in L^* .$$

Since

$$\mu^*\left(\frac{t-u}{\varepsilon}\right) f^*(t) \in I^*_{\mathfrak{M}},$$

the proof will be complete if we show that for sufficiently small $\varepsilon > 0$

$$\frac{h^*(t)}{1 + h^*(t)} \in L^* \ .$$

Let us choose $\varepsilon > 0$ so small that the inequality

$$|h^*(t)| < \frac{1}{2}$$

holds for all values of t. Then the function

$$\Phi(z) = \frac{z}{z+1}$$

will be regular on the curve $z = h^*(t)$ and equal to zero when $z = 0$. Therefore, by the Wiener-Levy theorem,

$$\frac{h^*(t)}{1 + h^*(t)} = \Phi[h^*(t)] \in L^*,$$

and the proof is complete.

Let ϱ be the radius of convergence of this series. In accordance with the statement (1) in Lemma 1, determine an $a > 0$ so that

$$\delta = \left\| f^*(t) - f^*(t)\, \mu^*\left(\frac{t}{a}\right) \right\| < \varrho$$

and set

$$g^*(t) = f^*(t) - f^*(t)\, \mu^*\left(\frac{t}{a}\right).$$

Then we have

$$\left\| \sum_{k=1}^{\infty} a_k [g^*(t)]^k \right\| \leqq \sum_{k=1}^{\infty} |a_k| \delta^k < \infty$$

and

$$\omega_\infty(t) = \Phi[g^*(t)] \in L^*.$$

Whence it obviously follows that equation (2) also is valid for $|t| \geqq 2a$.

From Lemma 2 we have

$$\Phi[f^*(t)] \in L^*,$$

and the theorem is proved.

102. Proof of the Sufficiency of Wiener's Condition. Let Wiener's condition be fulfilled, that is to say, suppose there does not exist any finite point u at which the Fourier transforms of all the functions in $\mathfrak{M}$ vanish. We will prove that an arbitrary function in L belongs to the ideal $I_{\mathfrak{M}}$, in other words, that an arbitrary function on L^* lies within the ideal $I^*_{\mathfrak{M}}$.

It suffices to show that any function $g^*(t) \in L^*$ which vanishes outside of a certain interval (i.e. a finite interval, dependent on the function selected), is contained in the ideal $I^*_{\mathfrak{M}}$. As a matter of fact, the set of these functions lies dense in L^*, by reason of the statement (2) in Lemma 1 of Section 100.

Now, suppose that $g^*(t) \in L^*$, and that $g^*(t) = 0$ when $|t| > c > 0$. Thus we have a certain neighborhood of the point at infinity where

$$g^*(t) = \omega_\infty(t) \equiv 0 \in I^*_{\mathfrak{M}}\ .$$

In order to be able to apply Lemma 2 of Section 100, it suffices to show that, by reason of Wiener's condition, for every finite point there exists likewise an $\varepsilon > 0$ such that

$$\mu^*\left(\frac{t-u}{\varepsilon}\right) \in I^*_{\mathfrak{M}}\ .$$

It then follows from the foregoing that

$$g^*(t)\, \mu^*\left(\frac{t-u}{\varepsilon}\right) \equiv \omega_u(t) \in I^*_{\mathfrak{M}}$$

and this function coincides with $g^*(t)$ when $u - \varepsilon \leqq t \leqq u + \varepsilon$.

103. Wiener's General Tauber Theorem. Wiener has employed his approximation theorem for the purpose of deriving a number of important interrelated propositions. These he calls Tauber theorems.

It is impossible to reproduce in detail in this book Wiener's investigations relevant to this subject. We will confine ourselves to the following theorem: *If* $\mathfrak{M} \in L$ *and there does not exist any finite point* u *at which the Fourier transforms of all the functions within* $\mathfrak{M}$ *vanish, and if for a certain bounded measurable function* $a(s)$ $(-\infty < s < \infty)$ *the equation*

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} f(T-s)\, a(s)\, ds = A \int_{-\infty}^{\infty} f(t)\, dt$$

is valid, where $f(s)$ *is an arbitrary function of the set* $\mathfrak{M}$ *and* A *is a certain constant, then we have*

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} g(T-s)\, a(s)\, ds = A \int_{-\infty}^{\infty} g(t)\, dt$$

for every function $g(s) \in L$.

Proof: If we preassign any positive number ε, then, by Wiener's approximation theorem, functions $f_i \in \mathfrak{M}$ and constants c_{ik}, λ_{ik} can be found such that

$$\int_{-\infty}^{\infty} \Big| g(s) - \sum_{i,k} c_{i,k} f_i(s+\lambda_{i,k}) \Big|\, ds < \varepsilon \ .$$

Hence it follows that

$$\Big| \int_{-\infty}^{\infty} \Big\{ g(t) - \sum_{i,k} c_{i,k} f_i(t+\lambda_{i,k}) \Big\}\, dt \Big| < \varepsilon \tag{1}$$

and

$$\Big| \int_{-\infty}^{\infty} g(T-s)\, a(s)\, ds - \sum_{i,k} c_{i,k} \int_{-\infty}^{\infty} f_i(T-s+\lambda_{i,k})\, a(s)\, ds \Big| \leqq$$

$$\leqq \operatorname*{true\,max}_{-\infty<s<\infty} |a(s)| \int_{-\infty}^{\infty} \Big| g(T-s) - \sum_{i,k} c_{i,k} f_i(T-s+\lambda_{i,k}) \Big|\, ds \leqq \varepsilon \operatorname*{true\,max}_{-\infty<s<\infty} |a(s)| .$$

If, in the last equation, we allow $T \to \infty$, we find:

$$\overline{\lim_{T\to\infty}} \Big| \int_{-\infty}^{\infty} g(T-s)\, a(s)\, ds - A \sum_{i,k} c_{i,k} \int_{-\infty}^{\infty} f_i(t+\lambda_{i,k})\, dt \Big| \leqq \varepsilon \operatorname*{true\,max}_{-\infty<s<\infty} |a(s)| .$$

Therefore, because of (1), we get:

$$\overline{\lim_{T\to\infty}} \left| \int_{-\infty}^{\infty} g(T-s)\, a(s)\, ds - A \int_{-\infty}^{\infty} g(t)\, dt \right| \leqq \varepsilon \{\text{true} \max_{-\infty<s<\infty} |a(s)| + |A|\}.$$

But since $\varepsilon > 0$ was arbitrarily chosen, the left-hand side is equal to zero, as we were to prove.

104. Weakly Decreasing Functions [83]. We say that the (real-valued) function $a(s)$ $(-\infty < s < \infty)$ for $s \to \infty$ decreases weakly if the following inequality holds:

$$\lim_{\substack{-s\to 0 \\ t>s,\, s\to\infty}} \{a(t) - a(s)\} \geqq 0.$$

Lemma: *If $a(s)$ for $s \to \infty$ decreases weakly and if*

$$\lim_{s\to\infty} a(s) > \delta > 0 \qquad (\text{analogously } \lim_{s\to\infty} a(s) < -\delta < 0) \tag{1}$$

then there exists an infinite sequence of intervals $E_n = (x_n, x_n + \varepsilon)$ $(x_n \to \infty)$ of constant length ε such that

$$a(s) > \frac{\delta}{2} \quad \left(\text{analogously } a(s) < -\frac{\delta}{2}\right)$$

in these intervals.

Proof: Since $a(s)$ for $s \to \infty$ decreases weakly, there exist two positive numbers N and ε such that for $s > N$ and $s < t < s + \varepsilon$ we have

$$a(t) - a(s) > -\frac{\delta}{4} \quad . \tag{2}$$

On the other hand, because of (1), there exists an infinite sequence $x_n (x_n \to \infty)$ for which the inequality

$$a(x_n) > \frac{3}{4}\delta$$

is valid. Obviously, we may further assume that $x_n > N$ $(n = 1, 2, \ldots)$. Hence, if we set $s = x_n$ in the relation (2) we find that the inequality

$$a(t) > a(x_n) - \frac{\delta}{4} > \frac{\delta}{2}$$

holds in the intervals $(x_n, x_n + \varepsilon)$.

The second statement in the lemma can be verified in an analogous way.

THEOREM: *Let $a(s)$ $(-\infty < s < \infty)$ be a bounded real-valued function which decreases weakly for $s \to \infty$ and let $\mathfrak{M} \subset L$. If there does not exist any finite point*

u at which the Fourier transforms of all the functions within $\mathfrak{M}$ *vanish, and if*

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} f(T-s)\, a(s)\, ds = A \int_{-\infty}^{\infty} f(t)\, dt$$

for every function $f \in \mathfrak{M}$, *then we have*

$$\lim_{T\to\infty} a(T) = A\,.$$

Proof: From the Tauber theorem in Section 103, we have

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} g(T-s)\, a(s)\, ds = A \int_{-\infty}^{\infty} g(t)dt$$

for any function $g(s) \in L$. Let us select as $g(s)$ the function

$$g_\lambda(s) = \lambda h(\lambda s)\,,$$

where

$$h(s) = \begin{cases} 1, & 0 \leqq s \leqq 1\,, \\ 0, & \left| s - \frac{1}{2} \right| > \frac{1}{2} \end{cases}$$

and λ is an arbitrary positive number. Hence we have for an arbitrary positive λ

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} \lambda h(\lambda T - \lambda s)\, a(s)\, ds = A$$

or

$$\lim_{T\to\infty} \lambda \int_{T}^{T+\frac{1}{\lambda}} \{a(s) - A\}\, ds = 0\,. \tag{3}$$

We are to prove:

$$\lim_{T\to\infty} b(T) = \lim_{T\to\infty} \{a(T) - A\} = 0\,.$$

Since $b(s)$ also decreases weakly we can again apply the lemma. Now, let us assume that

$$\overline{\lim_{s\to\infty}}\, b(s) > \delta > 0\ .$$

According to the lemma, there exist intervals $E_n = (x_n,\ x_n + \varepsilon)$ $(x_n \to \infty)$ in which $b(s) > \frac{\delta}{2}$. Let $\lambda = \frac{1}{\varepsilon}$. Then

$$\frac{1}{\varepsilon}\int_{x_n}^{x_n+\varepsilon} b(s)\,ds > \frac{\delta}{2} \qquad (n = 1, 2, 3, \ldots),$$

which leads to a contradiction in relation to (3).

Consequently we have

$$\lim_{s\to\infty} b(s) \leqq 0.$$

Analogously it can be verified that

$$\lim_{s\to\overline{\infty}} b(s) \geqq 0,$$

and the proof is complete.

105. Remarks on the Terminology. The designations "Tauber theorem" and "Tauber's asymptotic," which is often paired with "Abel's asymptotic," are due to Hardy and Littlewood. The reason for this terminology is that this typical way of formulating the problem appeared for the first time in the theory of power series in connection with the well-known theorems of Abel and Tauber.

Abel's theorem can be stated as follows: From

$$\lim_{n\to\infty} s_n = s$$

follows

$$\lim_{x\to 1-0} \sum_{n=0}^{\infty} (s_n - s_{n-1})\, x^n = s \qquad (s_{-1} = 0).$$

The converse of this theorem is untrue, as can be shown by simple examples. Nevertheless the following theorem due to Tauber is valid [84]: From

$$\lim_{x\to 1-0} \sum_{n=0}^{\infty} (s_n - s_{n-1}) x^n = s \qquad (s_{-1} = 0)$$

and

$$\lim_{n\to\infty} n(s_n - s_{n-1}) = 0$$

follows

$$\lim_{n\to\infty} s_n = s.$$

Here we have an operator connected with the sequence $\{s_n\}$ of the function

$$f(x) = \sum_{n=0}^{\infty} (s_n - s_{n-1})\, x^n \qquad (|x| < 1)$$

and our problem is to draw conclusions about the behavior of the sequence $\{s_n\}$ (the primitive) from the behavior of the function $f(x)$ (the image) as $x \to 1 - 0$

This circumstance makes it clear that Tauber's theorem and Wiener's theorem (especially the variant to which Section 104 is devoted) are theorems of the same kind.

106. Ikehara's Theorem [85]. As an application of Wiener's Tauber theorem we will prove the following theorem due to Ikehara: *Let* $\varphi(x)$ $(0 \leq x \leq \infty)$ *be a non-negative, non-decreasing function. If the Stieltjes integral*

$$f(s) = \int_0^\infty e^{-su}\, d\varphi(u) \quad (s = \sigma + i\tau)$$

is convergent for $\sigma > 1$ *and there exists a constant* A *such that the difference*

$$g(s) = f(s) - \frac{A}{s-1}$$

approaches a finite limit uniformly in every finite interval of the straight line $\sigma = 1$ *as* $\sigma \to 1 + 0$, then

$$\lim_{T\to\infty} \frac{\varphi(T)}{e^T} = A\,.$$

Proof: First of all we observe that

$$f(s) = s\int_0^\infty e^{-su}\varphi(u)\, du\ .$$

In fact, for $N > 0$ and $\sigma > 1$ we have

$$\int_0^N e^{-\sigma u}\, d\varphi(u) = e^{-\sigma N}\varphi(N) - \varphi(0) + \sigma\int_0^N e^{-\sigma u}\varphi(u)\, du\,,$$

whence it is evident that

$$\int_0^\infty e^{-\sigma u}\varphi(u)\, du$$

as well as

$$\lim_{N\to\infty} e^{-\sigma N}\varphi(N)$$

exists. This last limit can only be equal to zero because otherwise the preceding integral would be divergent.

In view of the above remark, we can start from the very beginning with the integral

$$F(s) = \int_0^\infty e^{-su}\varphi(u)\, du \tag{1}$$

and the difference

$$G(s) = F(s) - \frac{A}{s-1}. \tag{2}$$

Let T be a real-valued parameter, let n be a positive integer and let $\varepsilon > 0$. By recalling the formulas connected with the Fejér kernel, we readily obtain from (1) and (2) the following relation:

$$\frac{1}{2\pi}\int_{-n}^{n}\left(1-\frac{|v|}{n}\right)e^{ivT}G(iv+1+\varepsilon)\,dv = \frac{2}{\pi n}\int_0^\infty \varphi(u)\,e^{-u(1+\varepsilon)}\frac{\sin^2\frac{n}{2}(T-u)}{(T-u)^2}\,du -$$

$$-\frac{2A}{\pi n}\int_0^\infty e^{-\varepsilon u}\frac{\sin^2\frac{n}{2}(T-u)}{(T-u)^2}\,du.$$

In this relation let us pass to the limit by allowing $\varepsilon \to 0$.

It is easy to see that this passage to the limit under the integral sign can be accomplished. In the second integral of the right-hand side, that is obvious; on the left, it follows by hypothesis; and, finally, it follows in the first term on the right-hand side because the expression under the integral sign does not decrease as $\varepsilon \to 0$. Consequently,

$$\frac{1}{2\pi}\int_{-n}^{n}\left(1-\frac{|v|}{n}\right)e^{ivT}G(iv+1)\,dv =$$

$$= \frac{2}{\pi n}\int_0^\infty \varphi(u)\,e^{-u}\frac{\sin^2\frac{n}{2}(T-u)}{(T-u)^2}\,du - \frac{A}{\pi}\int_{-\infty}^{\frac{nT}{2}}\frac{\sin^2 v}{v^2}\,dv.$$

By virtue of the Riemann-Lebesgue theorem, the left-hand side approaches zero as $T \to \infty$. Therefore,

$$\lim_{T\to\infty}\frac{2}{\pi n}\int_0^\infty \frac{\sin^2\frac{n}{2}(T-u)}{(T-u)^2}\,\varphi(u)\,e^{-u}\,du = A \qquad (n = 1, 2, 3, \ldots). \tag{3}$$

We will prove that, by reason of this relation, the function

$$a(s) = \begin{cases} \varphi(s)\,e^{-s} & (s \geqq 0), \\ 0 & (s < 0) \end{cases}$$

is bounded and for $s \to \infty$ decreases weakly.

In fact, because of (3), there exists for $n = 1$ a T_0 such that, for $T > T_0$,

$$\frac{1}{\pi}\int_{-\infty}^{\infty} a(T+2v)\frac{\sin^2 v}{v^2}\,dv < A + 1$$

and therefore a fortiori

$$\frac{1}{\pi}\int_0^1 a(T+2v)\frac{\sin^2 v}{v^2}\,dv < A+1 \qquad (T > T_0).$$

But since

$$a(s_1)\,e^{s_1} \leqq a(s_2)\,e^{s_2} \qquad (s_1 \leqq s_2)$$

it follows that

$$\frac{1}{\pi}a(T)\int_0^1 e^{-2v}\frac{\sin^2 v}{v^2}\,dv < A+1 \qquad (T > T_0).$$

Hence, if $T > T_0$,

$$a(T) < \frac{\pi(A+1)\,e^2}{\int_0^1 \frac{\sin^2 v}{v^2}\,dv}.$$

Consequently $a(s)$ is bounded. On the other hand, if $t > s$, we have

$$a(t) - a(s) = \varphi(t)\,e^{-t} - \varphi(s)\,e^{-s} \geqq \varphi(s)\,(e^{-t} - e^{-s}) = a(s)\,(e^{s-t} - 1).$$

The second factor on the right-hand side approaches zero with $t - s$, but the first factor, as we have shown, is bounded. Whence it follows that $a(s)$ decreases weakly.

It is now easy to complete the proof of Ikehara's theorem. The equation (3) has the form

$$\lim_{T\to\infty}\int_{-\infty}^{\infty} f_n(T-s)\,a(s)\,ds = A \qquad (n = 1, 2, \ldots)$$

where

$$f_n(s) = \frac{2}{\pi n}\frac{\sin^2\frac{ns}{2}}{s^2}.$$

Here $a(s)$ is a bounded, weakly decreasing function and the sequence $\{f_n(s)\}_1^\infty$ constitutes a set $\mathfrak{M} \subset L$. Since

$$f_n^*(s) = \begin{cases} 1 - \dfrac{|s|}{n} & (|s| \leqq n), \\ 0 & (|s| > n) \end{cases}$$

there does not exist any point at which the Fourier transforms of all the functions within $\mathfrak{M}$ vanish. By applying the Tauber theorem of Section 104, we obtain

$$\lim_{T\to\infty} a(T) = A,$$

that is to say

$$\lim_{T\to\infty} \frac{\varphi(T)}{e^T} = A .$$

This proves Ikehara's theorem. It will be noted that this theorem plays an important role in analytical number theory. This theorem, in conjunction with the elementary properties of Riemann's zeta function, enables us to prove easily the famous approximation formula

$$\pi(x) \sim \frac{x}{\ln x}$$

for the number of primes not exceeding x, in the sequence of natural numbers.

107. Carleman's Tauber Theorem [86]: *Let* $\sigma(u)$ $(1 \leq u < \infty)$ *be a non-decreasing positive function. Furthermore, let the Stieltjes integral*

$$h(x) = \int_1^\infty \frac{d\sigma(u)}{u(u+x)}$$

exist for $x > 0$. *If we set*

$$h(x) = \frac{H \ln x}{x} + \frac{A}{x} + g(x), \tag{1}$$

where H *and* A *are constants and if*

$$\int_a^\infty |g(x)|\, dx < \infty \qquad (a > 0),$$

it follows that

$$\lim_{T\to\infty} \frac{\sigma(T)}{T} = H .$$

Proof: Let us take a complex number s with $1 < \Re(s) < 2$ and calculate the integral

$$\int_a^\infty \frac{h(x)}{x^{s-1}}\, dx$$

with $0 < a < 1$. It is easy to see that

$$\int_a^\infty \frac{h(x)}{x^{s-1}}\, dx = \int_1^\infty \frac{d\sigma(u)}{u} \left\{ \int_0^\infty \frac{dx}{(u+x)\, x^{s-1}} - \int_0^a \frac{dx}{(u+x)\, x^{s-1}} \right\} =$$

$$= \int_1^\infty \frac{d\sigma(u)}{u} \left\{ \frac{\pi}{u^{s-1} \sin \pi(s-1)} - \int_0^a \frac{dx}{(u+x)\, x^{s-1}} \right\} .$$

On the other hand, by integration along the path shown in Fig. 5 we find

$$\int_0^a \frac{dx}{(u+x)\,x^{s-1}} = \frac{a^{2-s}\,e^{\pi i s}}{2\sin\pi(s-1)} \int_0^{2\pi} \frac{e_{i\Theta}{}^{(2-s)}}{u + a e^{i\Theta}}\, d\Theta,$$

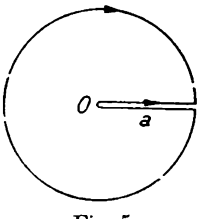


Fig. 5

and, from (1), it follows that

$$\int_a^\infty \frac{h(x)}{x^{s-1}}\,dx = \int_a^\infty \left\{ \frac{H\ln x}{x} + \frac{A}{x} + g(x) \right\} \frac{dx}{x^{s-1}} =$$

$$= \frac{Ha^{1-s}}{(s-1)^2} + \frac{a^{1-s}}{s-1}\{H\ln a + A\} + \int_a^\infty g(x)\,\frac{dx}{x^{s-1}}\,.$$

By reason of the given relations, we have

$$\int_1^\infty \frac{d\sigma(u)}{u^s} = \frac{H}{s-1} + G(s),$$

where $G(s)$ is a certain function which approaches a finite limit uniformly in every finite interval of the straight line $\Re(s) = 1$, as $\Re(s) \to 1$. Since

$$\int_1^\infty \frac{d\sigma(u)}{u^s} = \int_0^\infty e^{-st}\,d\sigma(e^t)$$

it follows from Ikehara's theorem that

$$\lim_{t\to\infty} \frac{\sigma(e^t)}{e^t} = H$$

or

$$\lim_{T\to\infty} \frac{\sigma(T)}{T} = H.$$

Thus Carleman's theorem has been proved. This theorem plays an essential role in a new method recently proposed by him, of determining the asymptotic behavior of eigenvalues of partial differential equations.

VARIOUS ADDENDA AND PROBLEMS

A. Elementary Extremal Problems and Certain Closure Criteria

1. Suppose that the points $z_1, z_2, \ldots, z_k$ lie in the region $|z| > 1$ and that the points $z_{k+1}, \ldots, z_n$ lie in the circle $|z| \leq 1$. Then, for $N \geq n$,

a) $$\min_{A_i} \max_{|z|=1} \frac{|z^N + A_1 z^{N-1} + \ldots + A_N|}{|(z-z_1)(z-z_2)\ldots(z-z_n)|} = \frac{1}{|z_1 z_2 \ldots z_k|},$$

b) for an arbitrary $p > 0$

$$\min_{A_i} \frac{1}{2\pi} \int_{|z|=1} \left| \frac{z^N + A_1 z^{N-1} + \ldots + A_N}{(z-z_1)(z-z_2)\ldots(z-z_n)} \right|^p |dz| = \frac{1}{|z_1 z_2 \ldots z_k|^p},$$

where the polynomial

$$z^N + A_1 z^{N-1} + \ldots + A_N,$$

for which the expression attains the minimum value, is always equal to

$$P(z) = \frac{z^{N-n}(z\bar{z}_1 - 1)\ldots(z\bar{z}_k - 1)(z - z_{k+1})\ldots(z - z_n)}{\bar{z}_1 \bar{z}_2 \ldots \bar{z}_k}$$

We are required to prove that, for any polynomial $P(z) \not\equiv Q(z)$ of the N-th degree with unity as the leading coefficient, the following equations hold:

a) $$\max_{|z|=1} \left| \frac{Q(z)}{\omega(z)} \right| > \max_{|z|=1} \left| \frac{P(z)}{\omega(z)} \right| = \frac{1}{|z_1 z_2 \ldots z_k|},$$

b) $$\frac{1}{2\pi} \int_{|z|=1} \left| \frac{Q(z)}{\omega(z)} \right|^p |dz| > \frac{1}{2\pi} \int_{|z|=1} \left| \frac{P(z)}{\omega(z)} \right|^p |dz| = \frac{1}{|z_1 z_1 \ldots z_k|^p} \qquad (p > 0)$$

with

$$\omega(z) = (z - z_1)(z - z_2)\ldots(z - z_n).$$

By reason of the inequality

$$\max_{|z|=1} \left| \frac{Q(z)}{\omega(z)} \right|^p \geqq \frac{1}{2\pi} \int_{|z|=1} \left| \frac{Q(z)}{\omega(z)} \right|^p |dz|$$

a) follows from b). To prove b) let us denote by $\alpha_1, \alpha_2, \ldots, \alpha_q$ the zeros of the polynomial $Q(z)$ which lie in the region $|z| < 1$, and then set

$$Q(z) = (z - \alpha_1)(z - \alpha_2)\ldots(z - \alpha_q) S(z),$$

where

$$|S(0)| \geqq 1 .$$

The function

$$R(z) = \frac{(1 - \bar{a}_1 z)(1 - \bar{a}_2 z) \ldots (1 - \bar{a}_q z) S(z)}{(z - z_1)(z - z_2) \ldots (z - z_k)(1 - \bar{z}_{k+1} z) \ldots (1 - \bar{z})}$$

is regular and different from zero in the region $|z| < 1$. Evidently we may assume that this function is continuous in the circle $|z| \leq 1$. These properties belong also to the function $[R(z)]^p$. Consequently

$$\int_{z=1} [R(z)]^p \frac{dz}{z} = 2\pi i [R(0)] .$$

But since the equation

$$|R(z)| = \left| \frac{Q(z)}{\omega(z)} \right|$$

is valid for $|z| = 1$, we have

$$\frac{1}{2\pi} \int_{z=1} \left| \frac{Q(z)}{\omega(z)} \right|^p |dz| = \frac{1}{2\pi} \int_{z=1} |R(z)|^p |dz| \geqq \frac{1}{2\pi} \left| \int_{z=1} [R(z)]^p \frac{dz}{z} \right| =$$

$$= |R(0)|^p = \frac{|S(0)|^p}{|z_1 z_2 \ldots z_k|^p} \geqq \frac{1}{|z_1 z_2 \ldots z_k|^p} .$$

It is easy to verify that the equality sign in these two relations cannot remain at all if $Q(z) \not\equiv P(z)$.

2. Suppose we have a sequence of numbers z_k $(k = 1, 2, 3, \ldots)$. Let the numbers of this sequence be distinct and satisfy the condition $|z_k| \neq 1$. We are to prove that the sequence of functions

$$\left\{ \frac{1}{e^t - z_k} \right\}_{k=1}^{\infty} \qquad (0 \leqq t \leqq 2\pi)$$

is closed in the space $L^p(0, 2\pi)$ $(p \geq 1)$ as well as in the space of continuous functions with the period 2π if and only if the equations

$$\sum_{k=1}^{\infty} \{1 - |z_k|^*\} = \infty, \quad \sum_{k=1}^{\infty} \left\{ 1 - \left| \frac{1}{z_k} \right|^* \right\} = \infty$$

hold with

$$|z|^* = \begin{cases} |z| & (|z| \leqq 1), \\ 1 & (|z| \geqq 1). \end{cases}$$

It suffices to show that the closure of this sequence of functions, in the space of all continuous periodic functions, is a consequence of the fulfillment of the conditions set in the problem, and that these conditions are necessary for the closure in the space $L(0, 2\pi)$.

Let the sequence of functions

$$\left\{ \frac{1}{e^{it} - z_k} \right\}_{k=1}^{\infty}$$

be closed in $L(0, 2\pi)$. Then the expression (see the previous section)

$$\min_{A_k} \int_0^{2\pi} \left| 1 - \sum_{k=1}^{n} \frac{A_k}{e^{it} - z_k} \right| dt = \min_{B_k} \int_{|z|=1} \left| \frac{z^n + B_1 z^{n-1} + \ldots + B_n}{(z - z_1) \ldots (z - z_n)} \right| |dz| = 2\pi \prod_{k=1}^{n} \left| \frac{1}{z_k} \right|^*$$

must approach zero as $n \to \infty$, and consequently the series

$$\sum_{k=1}^{\infty} \left\{ 1 - \left| \frac{1}{z_k} \right|^* \right\} \tag{1}$$

must be divergent.

Without loss of generality we may assume $z_k \neq 0$ $(k = 1, 2, 3, \ldots)$. Then it follows that

$$\min_{A_k} \int_0^{2\pi} \left| e^{-it} - A_0 - \sum_{k=1}^{n} \frac{A_k}{e^{it} - z_k} \right| dt =$$

$$= \min_{B_k} \int_{|z|=1} \left| \frac{z^{n+1} + B_1 z^n + \ldots + B_{n+1}}{\left(z - \frac{1}{\bar{z}_1}\right) \ldots \left(z - \frac{1}{\bar{z}_n}\right)} \right| |dz| = 2\pi \prod_{k=1}^{n} |z_k|^*.$$

Since this expression approaches zero as $n \to \infty$, the series

$$\sum_{k=1}^{\infty} \{ 1 - |z_k|^* \} \tag{2}$$

must also diverge.

Let us now assume that the two given series are divergent. From proposition a) of the theorem in Section 1, it follows that

$$\min_{A_k} \max_{|z|=1} \left| z^m - A_1 z^{m-1} - \ldots - A_m - \frac{A_{m+1}}{z - z_1} - \ldots - \frac{A_{m+n}}{z - z_n} \right| = \prod_{k=1}^{n} \left| \frac{1}{z_k} \right|^*$$

$$(m = 0, 1, 2, \ldots).$$

If we again assume $z_k \neq 0$ $(k = 1, 2, 3, \ldots)$, then we have

$$\min_{A_k} \max_{|z|=1} \left| \frac{1}{z^\mu} - \frac{A_1}{z^{\mu-1}} - \ldots - A_\mu - \frac{A_{\mu+1}}{z - z_1} - \ldots - \frac{A_{\mu+n}}{z - z_n} \right| =$$

$$= \min_{B_k} \max_{|z|=1} \left| z^\mu - B_1 z^{\mu-1} - \ldots - B_\mu - \frac{B_{\mu+1}}{z - \frac{1}{\bar{z}_1}} - \ldots - \frac{B_{\mu+n}}{z - \frac{1}{\bar{z}_n}} \right| =$$

$$= \prod_{k=1}^{n} |z_k|^* \qquad (\mu = 1, 2, \ldots).$$

Hence, if we set successively $m = 0, 1, 2, \ldots$; $n = 1, 2, 3, \ldots$, we find that each of the functions

$$1, e^{\pm it}, e^{\pm 2it}, \ldots \tag{3}$$

can be approximated with any degree of accuracy through combinations of given functions. We may assume without restriction (Weierstrass' second theorem) that set (3) is closed in the space of periodic continuous functions.

Series (1) and (2) are encountered continually in the modern theory of analytic functions (in relation to Blaschke's theorem and its generalizations). Therefore, from the standpoint of function theory it is no accident that these series appear in the problem under consideration. The reader who is familiar with this branch of function theory will have no difficulty in arriving at a similar proof of the criterion in question with the help of the theorems in Sections 29, 30.

3. Let $c_0, c_1, \ldots, c_n$ be complex numbers. It is required to prove that the following equations are valid:

a)
$$\min_{A_k} \max_{-\infty<x<\infty} \left| \frac{1}{x-c_0} - A_0 - \frac{A_1}{x-c_1} - \ldots - \frac{A_n}{x-c_n} \right| = \frac{1}{|c_0-\bar{c}_0|} \prod_{k=1}^{n} \left| \frac{c_0-c_k}{c_0-\bar{c}_k} \right|^*,$$

b)
$$\min_{A_k} \int_{-\infty}^{\infty} \left| \frac{1}{x-c_0} - \frac{A_1}{x-c_1} - \ldots - \frac{A_n}{x-c_n} \right|^2 dx = \frac{2\pi}{|c_0-\bar{c}_0|} \left\{ \prod_{k=1}^{n} \left| \frac{c_0-c_k}{c_0-\bar{c}_k} \right|^* \right\}^2,$$

c) for an arbitrary $p > 0$

$$\min_{A_k} \int_{-\infty}^{\infty} \left| \frac{1}{x-c_0} - A_0 - \frac{A_1}{x-c_1} - \ldots - \frac{A_n}{x-c_n} \right|^p \frac{dx}{|x-c_0|^2} = \frac{2\pi}{|c_0-\bar{c}_0|^{p+1}} \left\{ \prod_{k=1}^{n} \left| \frac{c_0-c_k}{c_0-\bar{c}_k} \right|^* \right\}^p.$$

The simplest method of proof is to make use of the conformal mapping of a half-plane upon a circle

$$z = \frac{x-\bar{c}_0}{x-c_0},$$

by means of which the problem is reduced to the problem treated in Section **1**.

4. Given a sequence of numbers c_k $(k = 1, 2, \ldots)$. Let no two of these numbers be identical, and suppose the inequality $\Im c_k \neq 0$ is satisfied.

It is required to prove that the sequence of functions

$$1, \quad \frac{1}{x-c_1}, \quad \frac{1}{x-c_2}, \ldots$$

is closed in the space C_∞ if and only if the two series

$$\sum' \frac{\Im c_k}{1+|c_k|^2}, \quad \sum'' \frac{\Im c_k}{1+|c_k|^2}$$

are divergent. Here the summation in $\sum'$ is to be extended over those values of k for which $\Im c_k > 0$ and in $\sum''$ over the remaining values of k. The same condition is necessary and sufficient for the closure of the sequence of functions

$$\left\{\frac{1}{x-c_k}\right\}_{k=1}^{\infty}$$

in the space $L^p(-\infty, \infty)$ for any $p > 1$.

In an abstruse form, this theorem has been given by Bernstein [68].

First of all, we observe merely that the divergence of the aforementioned series is a necessary and sufficient condition that

$$\lim_{n\to\infty} \prod_{n=1}^{n} \left|\frac{c-c_k}{c-\bar{c}_k}\right|^* = 0$$

for any complex value of c.

In order to prove this statement, let us assume for the sake of definiteness that

$$\Im c > 0$$

where naturally $c \neq c_k$ $(k = 1, 2, \ldots)$. Let us consider the product

$$\prod \left|\frac{c-c_\alpha}{c-\bar{c}_\alpha}\right|;$$

here α runs through those values of $k \leq n$ for which $\Im c_k > 0$. This product approaches zero if and only if

$$\sum' \left\{1 - \left|\frac{c-c_k}{c-\bar{c}_k}\right|^2\right\} = \infty \ .$$

Finally we observe that, for $\Im c > 0$, there exist [1] such $m = m(c) > 0$, $M = M(c) > m(c)$ that

[1] E.g. we can set

$$m = 4\frac{\Im c}{1+|c|^2}, \quad M = 8\frac{1+|c|^2}{\Im c} \ .$$

$$m\frac{\Im\gamma}{1+|\gamma|^2} \leqq 1-\left|\frac{c-\gamma}{c-\bar{\gamma}}\right|^2 \leqq M\frac{\Im\gamma}{1+|\gamma|^2}$$

follows from $\Im\gamma > 0$.

Let us now return to our problem and consider first the space C_∞. If $f(x)$ belongs to C_∞, then $g(t) = f\left(\operatorname{tg}\frac{t}{2}\right)$ is a function which is continuous in the interval $[-\pi, \pi]$ and such that $g(-\pi) = g(\pi)$. By means of the transformation

$$x = \operatorname{tg}\frac{t}{2}$$

the function

$$\frac{1-ix}{x-c_k}$$

is therefore carried into the function

$$\text{const.}\frac{1}{e^{it}-z_k} \qquad \left(z_k = \frac{1+ic_k}{1-ic_k}\right)$$

and our problem becomes that considered in Section 2, where the necessary and sufficient condition for the closure proved to be

$$\sum_{k=1}^{\infty}\{1-|z_k|^*\} = \sum_{k=1}^{\infty}\left\{1-\left|\frac{i-c_k}{i-\bar{c}_k}\right|^*\right\} = \infty, \quad \sum_{k=1}^{\infty}\left\{1-\left|\frac{i+c_k}{i+\bar{c}_k}\right|^*\right\} = \infty\ .$$

By reason of the remark made previously, this condition is equivalent to the one that appears in the statement of the theorem.

Let us now pass to the space $L^p(-\infty, \infty)$ $(p > 1)$ and prove the sufficiency of the condition. Let $f(x) \in L^p$. Corresponding to every $\varepsilon > 0$, there exists a continuous function $f_\varepsilon(x)$ equal to zero outside that interval for which

$$\int_{-\infty}^{\infty}|f(x)-f_\varepsilon(x)|^p\,dx \leqq \varepsilon^p\,. \tag{4}$$

Let us choose a point c_1 of our sequence and consider the function

$$g_\varepsilon(x) = (x-c_1)\,f_\varepsilon(x)\,,$$

which likewise belongs to C_∞. Then there exists an A_k with

$$\max_{-\infty \leqq x \leqq \infty}\left|g_\varepsilon(x) - A_1 - \sum_{k=2}^{N}\frac{A_k}{x-c_k}\right| \leqq \varepsilon\,,$$

and consequently we have

$$f(x) - \frac{A_1}{x-c_1} - \sum_{k=2}^{N}\frac{A_k}{(x-c_1)(x-c_k)}\Bigg| \leqq \frac{\varepsilon}{|x-c_1|} \qquad (-\infty \leqq x \leqq \infty)$$

or

$$f_\varepsilon(x) - \sum_{k=1}^{N} \frac{B_k}{x - c_k} \Big| \leqq \frac{\varepsilon}{|x - c_1|}.$$

Thus

$$\int_{-\infty}^{\infty} \Big| f_\varepsilon(x) - \sum_{k=1}^{N} \frac{B_k}{x - c_k} \Big|^p dx \leqq \varepsilon^p \int_{-\infty}^{\infty} \frac{dx}{|x - c_1|^p},$$

and, in conjunction with (4), this proves the sufficiency of the condition.

Now we come to the proof of the necessity of the condition: Let us consider the function $\frac{1}{x - c}$, where $\Im c \neq 0$ and $c \neq c_k$ lies in L^p. Suppose

$$\int_{-\infty}^{\infty} \Big| \frac{1}{x - c} - \sum_{k=1}^{n} \frac{A_k}{x - c_k} \Big|^p dx \leqq \varepsilon.$$

From

$$\int_{-\infty}^{\infty} \Big| \frac{1}{x - c} - \sum_{k=1}^{n} \frac{A_k}{x - c_k} \Big|^p \frac{dx}{|x - c|^2} \leqq \frac{4}{|c - \bar{c}|^2} \int_{-\infty}^{\infty} \Big| \frac{1}{x - c} - \sum_{k=1}^{n} \frac{A_k}{x - c_k} \Big|^p dx$$

it follows by virtue of proposition c) in Section 3 that

$$\frac{2\pi}{|c - \bar{c}|^{p+1}} \left\{ \prod_{k=1}^{n} \left| \frac{c - c_k}{c - \bar{c}_k} \right|^* \right\}^p \leqq \frac{4\varepsilon}{|c - \bar{c}|^2},$$

and it remains only to take into account our introductory remark.

5. Given a polynomial [2]

$$\omega(x) = \left(1 - \frac{x}{a_1}\right)\left(1 - \frac{x}{a_2}\right) \dots \left(1 - \frac{x}{a_{2q}}\right),$$

which is positive in the closed interval $[-1, 1]$. Let us set

$$x = \frac{1}{2}\left(v + \frac{1}{v}\right) \qquad (|v| \leqq 1),$$

$$a_k = \frac{1}{2}\left(c_k + \frac{1}{c_k}\right) \qquad (|c_k| < 1,\ k = 1, 2, \dots, 2q),$$

[2] The degree of $\omega(x)$ is odd $(2q - 1)$ if $a_{2q} = \infty$, $|a_k| < \infty$ $(k = 1, 2, \dots 2q - 1)$.

$$\Omega(v) = \prod_{k=1}^{2q} \sqrt{v - c_k},$$

$$\mathscr{L}_m = \begin{cases} \dfrac{1}{2^{m-1}} \prod\limits_{k=1}^{2q} \sqrt{1 + c_k^2} & (m > q), \\ \dfrac{1}{2^{q-1}} \dfrac{1}{1 + c_1 c_2 \ldots c_{2q}} \prod\limits_{k=1}^{2q} \sqrt{1 + c_k^2} & (m = q). \end{cases}$$

Then

$$T_m(x; \omega) = \frac{\mathscr{L}_m}{2} \left\{ v^{2q-m} \frac{\Omega\left(\frac{1}{v}\right)}{\Omega(v)} + v^{m-2q} \frac{\Omega(v)}{\Omega\left(\frac{1}{v}\right)} \right\} \sqrt{\omega(x)}$$

(where the positive integer $m \geq q$) is a polynomial of degree m in x, with unity as the leading coefficient.

It is required to prove:

a) $$\min_{A_k} \max_{-1 \leq x \leq 1} \frac{|x^m + A_1 x^{m-1} + \ldots + A_m|}{\sqrt{\omega(x)}} = \mathscr{L}_m,$$

where $T_m(x, \omega)$ is the extremal polynomial; [3]

b) $$\int_{-1}^{1} T_m(x; \omega) \frac{x^k}{\omega(x)\sqrt{1 - x^2}} dx = \begin{cases} 0 & (k = 0, 1, 2, \ldots, m-1), \\ \pi \mathscr{L}_m \mathscr{L}_{m+1} & (k = m) \end{cases}$$

$$(m = q, q+1, \ldots)$$

with

$$\prod_{k=1}^{2q} (1 + c_k^2) = e^{-\frac{1}{\pi} \int_{-1}^{1} \frac{\ln \omega(t)}{\sqrt{1-t^2}} dt}$$

The formulation of the theorem, as given in this and in the following Section, is due to Bernstein [87]. A special case thereof was considered earlier [88].

In order to prove proposition a), it suffices to make use of Tchebysheff's general theorem in Section 34 and carry out the change in argument of the function

$$v^{2q-m} \frac{\Omega\left(\frac{1}{v}\right)}{\Omega(v)}$$

if the point v describes the upper half of the circle $|v| = 1$, similar to what was done in Section 37.

[3] "Extremal polynomial" in this sense always means the polynomial which produces an extremum. — Trans.

The proof of proposition b) can be derived from the relation

$$\int_{-1}^{1} T_m(x;\omega)\frac{x_k}{\omega(x)\sqrt{1-x^2}}dx =$$

$$= \frac{\mathscr{L}_m}{2^{k+1}}\prod_{r=1}^{2q}\sqrt{1+c_r^2}\int_0^{\pi}\left\{v^{2q-m}\frac{\Omega\left(\frac{1}{v}\right)}{\Omega(v)} + v^{m-2q}\frac{\Omega(v)}{\Omega\left(\frac{1}{v}\right)}\right\}\times\left(v+\frac{1}{v}\right)^k\frac{d\vartheta}{\Omega(v)\,\Omega\left(\frac{1}{v}\right)} =$$

$$= \frac{\mathscr{L}_m\mathscr{L}_{m+1}}{2^{k-m+1}}\int_{v=1} v^{m-2q}\frac{\left(v+\frac{1}{v}\right)^k}{\left[\Omega\left(\frac{1}{v}\right)\right]^2}\frac{dv}{iv} \qquad (v=e^{i\vartheta})$$

by a simple application of the theorem on residues.

6. (Continuation). Let us weaken the hypothesis of the previous theorem and allow the polynomial $\omega(x)$, which is positive in the interior of the interval $(-1, 1)$, to have simple roots at one or both ends of the interval $(-1, 1)$.

Under the new hypothesis, we are required to prove:

a) $$U_m(x;\omega) = \mathscr{L}_{m+1}\left\{v^{2q-m-1}\frac{\Omega\left(\frac{1}{v}\right)}{\Omega(v)} - v^{m+1-2q}\frac{\Omega(v)}{\Omega\left(\frac{1}{v}\right)}\right\}\frac{\sqrt{\omega(x)}}{\frac{1}{v}-v}$$

is a polynomial of the m-th degree with unity as the leading coefficient. Here $m \geq q$, where m is a positive integer;

b) $$\int_{-1}^{1} U_m(x;\omega)\frac{x^k\sqrt{1-x^2}}{\omega(x)}dx = \begin{cases} 0 & (k=0,1,2,\ldots,m-1), \\ \frac{\pi}{2}\mathscr{L}_{m+1}^2 & (k=m) \end{cases}$$

$$(m = q, q+1, q+2, \ldots),$$

c) $$\min_{A_k}\int_{-1}^{1}\left|(x^m + A_1x^{m-1} + \ldots + A_m)\frac{\sqrt{1-x^2}}{\sqrt{\omega(x)}}\right|^p\frac{dx}{\sqrt{1-x^2}} =$$

$$= \int_{-1}^{1}\left|U_m(x;\omega)\frac{\sqrt{1-x^2}}{\sqrt{\omega(x)}}\right|^p\frac{dx}{\sqrt{1-x^2}} = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}+1\right)}\mathscr{L}_{m+1}^p$$

$$(m = q, q+1, q+2, \ldots)$$

when $p \geq 1$.

Proposition b) is proved exactly in the same way as the corresponding proposition in Section 5.

When $p = 2$, proposition c) follows from b).

Let us prove c) for $p = 1$. Since

$$v^{m+1-2q}\frac{\Omega(v)}{\Omega\left(\frac{1}{v}\right)} = e^{i\Phi}$$

for $v = e^{i\vartheta}$ and a real-valued Φ, it follows that

$$U_m(x;\omega) = \mathscr{L}_{m+1}\sqrt{\omega(x)}\cdot\frac{\sin\Phi}{\sin\vartheta} \qquad (-1 \leqq x \leqq 1)$$

and therefore

(1) $\operatorname{sign} U_m(x;\omega) = \operatorname{sign}\sin\Phi =$

$$= \frac{4}{\pi}\sum_{r=0}^{\infty}\frac{\sin(2r+1)\Phi}{2r+1} = \frac{2}{\pi i}\sum_{r=0}^{\infty}\frac{1}{2r+1}\left\{\left[v^{m+1-2q}\frac{\Omega(v)}{\Omega\left(\frac{1}{v}\right)}\right]^{2r+1}\right.$$

$$\left. - \left[v^{2q-m-1}\frac{\Omega\left(\frac{1}{v}\right)}{\Omega(v)}\right]^{2r+1}\right\} =$$

$$= \frac{2}{\pi i}\sum_{r=0}^{\infty}\frac{1}{(2r+1)\,\mathscr{L}_{m+1}^{2r+1}}U_{(2r+1)m}(x;\omega^{2r+1})\frac{v-\frac{1}{v}}{[\sqrt{\omega(x)}]^{2r+1}}.$$

Hence, according to proposition b), we have

$$\int_{-1}^{1}\operatorname{sign} U_m(x;\omega)\frac{x^k\,dx}{\sqrt{\omega(x)}} = \frac{2}{\pi i}\sum_{r=0}^{\infty}\frac{1}{(2r+1)\,\mathscr{L}_{m+1}^{2r+1}}\int_{-1}^{1}U_{(2r+1)m}(x;\omega^{2r+1})\,x^k\,[\omega(x)]^r \times$$

$$\times\frac{\sqrt{1-x^2}}{[\omega(x)]^{2r+1}}dx = 0 \qquad (k = 0, 1, 2, \ldots, m-1).$$

After obtaining this relation, we need only repeat almost literally the corresponding reasoning in Section 50.

We now give the proof due to Bernstein of proposition b) for $p > 1$. When $p > 2$, this proof is obtained quite readily by making use of the inequality

$$y^p - z^p \geqq \frac{p}{2}z^{p-2}(y^2 - z^2),$$

which is easy to verify and is valid for any $y > 0$, $z > 0$, $p > 2$.

In fact, if $Q(x)$ is any polynomial of the m-th degree with unity as the leading coefficient, then, by virtue of this inequality, we have

$$I = \int_{-1}^{1} \{ |Q(x)|^p - |U_m(x;\omega)|^p \} \left[\frac{1-x^2}{\omega(x)} \right]^{\frac{p}{2}} \frac{dx}{\sqrt{1-x^2}} \geqq$$

$$\geqq \frac{p}{2} \int_{-1}^{1} |U_m(x;\omega)|^{p-2} \{ Q^2(x) - U_m^2(x;\omega) \} \left[\frac{1-x^2}{\omega(x)} \right]^{\frac{p}{2}} \frac{dx}{\sqrt{1-x^2}} =$$

$$= \frac{p}{2} \int_{-1}^{1} \left| U_m(x;\omega) \frac{\sqrt{1-x^2}}{\omega(x)} \right|^{p-2} \{ Q^2(x) - U_m^2(x;\omega) \} \frac{\sqrt{1-x^2}}{\omega(x)} dx =$$

$$= \frac{p}{2} \mathcal{L}_{m+1}^{p-2} \int_{-1}^{1} |\sin \Phi|^{p-2} \{ Q^2(x) - U_m^2(x;\omega) \} \frac{\sqrt{1-x^2}}{\omega(x)} dx .$$

It is to be observed that

$$|\sin \Phi|^{p-2} = a_0 + a_2 \cos 2\Phi + a_4 \cos 4\Phi + \ldots$$

where $a_0 > 0$ and that

$$\cos 2r\Phi = \frac{1}{2} \left\{ \left[v^{m+1-2q} \frac{\Omega(v)}{\Omega\left(\frac{1}{v}\right)} \right]^{2r} + \left[v^{2q-m-1} \frac{\Omega\left(\frac{1}{v}\right)}{\Omega(v)} \right]^{2r} \right\} =$$

$$= \text{const.} \frac{T_{2r(m+1)}(x;\omega^{2r})}{[\omega(x)]^r}; \tag{2}$$

moreover, $Q^2(x) - U_m^2(x;\omega)$ is a polynomial of degree $2m-1$; and we obtain the relation

$$\int_{-1}^{1} \frac{T_{2r(m+1)}(x;\omega^{2r})}{[\omega(x)]^r} \{ Q^2(x) - U_m^2(x;\omega) \} \frac{\sqrt{1-x^2}}{\omega(x)} dx = 0$$

$$(r = 1, 2, 3, \ldots).$$

Consequently

$$I \geqq \frac{p}{2} \mathcal{L}_{m+1}^{p-2} a_0 \int_{-1}^{1} \{ Q^2(x) - U_m^2(x;\omega) \} \frac{\sqrt{1-x^2}}{\omega(x)} dx \geqq 0,$$

since the proposition has been proved for $p = 2$. Hence the extremal polynomial is equal to $U_m(x;\omega)$, that is, the same polynomial as in the cases $p = 1$ and $p = 2$.

In particular, by virtue of elementary properties of the extremum, it follows, from what has been established above, that for any $p \geq 2$ as well as for $p = 1$ we have

$$\int_{-1}^{1} \left| \frac{U_m(x;\omega)\sqrt{1-x^2}}{\sqrt{\omega(x)}} \right|^{p-1} \operatorname{sign} U_m(x;\omega) \frac{x^k dx}{\sqrt{\omega(x)}} = 0 \tag{3}$$

$$(k = 0, 1, 2 \ldots, m-1).$$

Again making use of the development (1), we can conclude that

$$\int_{-1}^{1} \cos 2r\Phi \operatorname{sign} U_m(x;\omega) \frac{x^k dx}{\sqrt{\omega(x)}} = 0$$
$$(k = 0, 1, \ldots, m-1)$$

for $r = 0, 1, 2, \ldots$ But

$$\left| \frac{U_m(x;\omega)\sqrt{1-x^2}}{\sqrt{\omega(x)}} \right| = a_0^{(\varrho)} + a_2^{(\varrho)} \cos 2\Phi + a_4^{(\varrho)} \cos 4\Phi + \ldots \tag{4}$$

for an arbitrary $\varrho > 0$, hence the relation (3) holds for every $p \geq 1$. Thus, for $p > 1$ the extremal polynomial, whose uniqueness results from the fact that the space is strictly normalized, is always $U_m(x;\omega)$.

The only thing left is to evaluate the integral

$$\mu_p = \int_{-1}^{1} \left| U_m(x;\omega) \frac{\sqrt{1-x^2}}{\sqrt{\omega(x)}} \right|^p \frac{dx}{\sqrt{1-x^2}} .$$

The development (4) and the relation (2) show that the equation

$$\mu_p = \int_{-1}^{1} a_0^{(\varrho)} \frac{dx}{\sqrt{1-x^2}} = \pi a_0^{\varrho} = \mathscr{L}_{m+1}^{p} \int_0^{\pi} \sin^p \Phi \, d\Phi$$

is valid, and therefore

$$\mu_p = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}+1\right)} \mathscr{L}_{m+1}^{p}.$$

7. Given a sequence of numbers $a_1, a_2, a_3, \ldots$. Suppose no two of these numbers are the same, and all of them lie in the interior of the interval $[-1, 1]$ on the real axis of the complex plane. It is required to prove that the sequence of functions

$$\left\{ \frac{1}{x - a_k} \right\}_{k=1}^{\infty} \tag{1}$$

is closed in the space C or in $L^p (p \geq 1)$, with respect to the interval $[-1, 1]$, if and only if

$$\sum_{k=1}^{\infty} \{1 - |c_k|\} = \infty$$

where

$$c_k = a_k - \sqrt{a_k^2 - 1} \qquad (|c_k| < 1,\ k = 1, 2, 3, \ldots).^4$$

[4] This theorem was imparted to the author orally by M. G. Krein. The proof given here differs from Krein's.

It suffices to show that the divergence of the series

$$\sum_{k=1}^{\infty}\{1-|c_k|\} \tag{2}$$

results in the closure of the sequence of functions (1) in the space C and that it is necessary for the closure in the space L.

Suppose the series (2) is divergent. Then, by virtue of a) of the theorem in Section **5**, we have for any integral value of $r \geq 0$:

$$\min_{A} \max_{-1\leq x\leq 1} \left| x^r - A_1 x^{r-1} - \ldots - A_r - \sum_{i=1}^{n} \frac{A_{r+i}}{x-a_i} \right| =$$

$$= \min_{B_k} \max_{-1\leq x\leq 1} \left| \frac{x^{n+r} + B_1 x^{n+r-1} + \ldots + B_{n+r}}{(x-a_1)(x-a_2)\ldots(x-a_n)} \right| \leqq$$

$$\leqq \frac{1}{2^{n+r-1}} \frac{\prod_{k=1}^{n} |1+c_k^2|}{|a_1 a_2 \ldots a_n|} = \frac{1}{2^{r-1}} \prod_{k=1}^{n} |c_k| \to 0 \quad (n\to\infty).$$

Thus we see that the sequence (1) is closed in the space of all polynomials and consequently in C.

Let us now assume that the sequence (1) is closed in L; then it follows that the quantity

$$I_n = \min_{A_k} \int_{-1}^{1} \left| 1 - \sum_{k=1}^{n} \frac{A_k}{x-a_k} \right| dx = \min_{B_k} \int_{-1}^{1} \left| \frac{x^n + B_1 x^{n-1} + \ldots + B_n}{(x-a_1)(x-a_2)\ldots(x-a_n)} \right| dx$$

approaches zero as $n \to \infty$.

But, by proposition c) in Section **6** we have (for $p = 1$)

$$I_n = \frac{1}{2^{n-1}} \frac{\prod_{k=1}^{n} |1+c_k^2|}{|a_1 a_2 \ldots a_n|} = 2 \prod_{k=1}^{n} |c_k|.$$

Whence the divergence of the series (2) follows.

8. Suppose the numbers

$$\sigma_k = a_k + i\beta_k \qquad (k = 0, 1, 2, \ldots, n)$$

fulfill the inequalities

$$|\beta_k| < \frac{\pi}{2} \qquad (k = 0, 1, 2, \ldots, n).$$

To prove:

$$\min_{A_k} \int_{-\infty}^{\infty} \left| e^{i\alpha_0 x} - \sum_{k=1}^{n} A_k e^{i\alpha_k x} \right|^2 \frac{dx}{\operatorname{ch} \pi x} =$$

$$= \frac{1}{\cos \beta_0} \prod_{k=1}^{n} \left\{ 1 - \frac{2 \cos \beta_0 \cos \beta_k}{\operatorname{ch} (a_0 - a_k) + \cos (\beta_0 + \beta_k)} \right\} .$$

Here the method of Section 14 can be used.

9. Given a sequence of numbers $\sigma_1, \sigma_2, \ldots$. Let no two numbers of this sequence be identical, and suppose the following inequalities fulfilled:

$$|\Im \sigma_n| < \frac{\pi}{2} \qquad (n = 1, 2, 3, \ldots) .$$

It is required to prove that the relation

$$\left\{ e^{-\frac{\pi}{2} x - i\sigma_n x} \right\}_{n=1}^{\infty}$$

is necessary and sufficient for the closure of the sequence of functions

$$\sum_{n=1}^{\infty} \frac{\cos \Im \sigma_n}{\operatorname{ch} \Re \sigma_n} = \infty$$

in $L^2(-\infty, \infty)$.

The theorem is contained in a book by Wiener and Paley [54]. The proof follows immediately from the results in Section 8 if we note that the sequence

$$\left\{ e^{-\frac{\pi}{2} x - i\sigma_n x} \right\}_{n=1}^{\infty}$$

is closed in $L^2(-\infty, \infty)$ if and only if the sequence

$$\left\{ \frac{e^{-i\sigma_n x}}{\sqrt{\operatorname{ch} \pi x}} \right\}_{n=1}^{\infty}$$

is closed.

B. Szegö's Theorem and Some of Its Applications

10. Let $w(t) \geq 0$ $(-\pi \leq t \leq \pi)$ be an integrable function such that

$$\int_{-\pi}^{\pi} w(t)\, dt > 0 .$$

We follow G. Szegö [89] and call the expression

$$\mathfrak{G}(w) = \begin{cases} \exp\left\{\frac{1}{\pi}\int_{-\pi}^{\pi} \ln w(t)\, dt\right\}, & \text{if} \quad \ln w(t) \in L, \\ 0, & \text{if} \quad \ln w(t) \overline{\in} L \end{cases}$$

the geometric mean of the function $w(t)$.

Szegö proved: For every $p > 0$, we have

$$\lim_{n\to\infty} \min_{A_k} \frac{1}{2\pi}\int_{-\pi}^{\pi} |z^n + A_1 z^{n-1} + \ldots + A_n|^p\, w(t) dt = \mathfrak{G}\{w\} \qquad (z = e^{it}).$$

By introducing the notation

$$\mu_n(w; p) = \min_{A_k} \frac{1}{2\pi}\int_{-\pi}^{\pi} |z^n + A_1 z^{n-1} + \ldots + A_n|^p\, w(t)\, dt$$

we find immediately

$$\mu_{n+1}(w; p) \leqq \mu_n(w; p).$$

Hence the existence of the limit

$$\lim_{n\to\infty} \mu_n(w; p) = \mu(w; p) \geqq 0$$

is ensured.

Let us assume first that $\mathfrak{G}\{w\} > 0$.

Then, by virtue of Jensen's inequality, we obtain for every polynomial

$$Q(z) = z^n + c_1 z^{n-1} + \ldots + c_n$$

the relation

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |Q(z)|^p\, w(t)\, dt \geqq \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} [\ln w(t) + \ln |Q(z)|^p]\, dt\right\} =$$

$$= \mathfrak{G}\{w\} \cdot \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln |Q(z)|^p\, dt\right\} = \mathfrak{G}\{w\}\, |z_1 z_1 \ldots z_k|^p,$$

where $z_1, z_2, \ldots, z_k$ are those roots of the polynomial $Q(z)$ which lie in the region $|z| > 1$.

From the inequality thus obtained, it follows that

$$\mu(w; p) \geqq \mathfrak{G}\{w\}.$$

We must therefore prove that the inequality

$$\mu(w; p) \leqq \mathfrak{G}\{w\}. \tag{1}$$

is valid.

First of all, we observe that if we set

$$w(t) = \frac{1}{|P(z)|^p},$$

where $P(z)$ is a polynomial, all the roots of which lie in the region $|z| > 1$, then from proposition b) in Section 1 we have, for every value of n sufficiently large,

$$\min_{A_k} \frac{1}{2\pi} \int_{-\pi}^{\pi} |z^n + A_1 z^{n-1} + \ldots + A_n|^p w(t)\, dt =$$

$$= \min_{A_k} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{z^n + A_1 z^{n-1} + \ldots + A_n}{P(z)} \right|^p dt = \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \frac{1}{|P(\epsilon^{it})|^p}\, dt \right\}.$$

In other words: In this case the equation

$$\mu(w; p) = \mathfrak{G}\{w\} \tag{2}$$

is surely valid. Now it is easy to show that (2) is valid if $w(t)$ is a continuous function satisfying the inequality $w(t) \geq \varrho > 0$. In fact, by Weierstrass' second theorem, we can find, in this case, non-negative trigonometric sums $S_m(t)$, $s_m(t)$ so that we have

$$\frac{1}{S_m(t)} < [w(t)]^{\frac{2}{p}} > \frac{1}{s_m(t)} \tag{3}$$

and the functions $\frac{1}{S_m(t)}$ and $\frac{1}{s_m(t)}$ approach $[w(t)]^{\frac{2}{p}}$ uniformly as $m \to \infty$. But since every non-negative trigonometric sum $Q(t)$ admits a representation

$$\sigma(t) = |P(z)|^2 \qquad (z = \epsilon^{it})$$

and since, on the other hand, we have from (3)

$$\mu\left(\frac{1}{S_m^{\frac{p}{2}}},\ p \right) \leqq \mu(w; p) \leqq \mu\left(\frac{1}{s_m^{\frac{p}{2}}};\ p \right)$$

it follows, from what has already been proved, that equation (2) is valid for every continuous function $w(t) \geq \varrho > 0$.

Now let us assume that $w(t)$ is any integrable function but retain the hypothesis $w(t) \geq \varrho > 0$.

By the theorem in Section 24, for every $\varepsilon > 0$ we can find a continuous function $f(t)$ so that

$$f(t) \geqq \frac{\varrho}{2}, \quad \int_{-\pi}^{\pi} |f(t) - w(t)|^2\, dt < \varepsilon^2 .$$

Then we obtain

$$\ln f(t) - \ln w(t) \leqq \frac{2}{\varrho} |f(t) - w(t)|$$

and accordingly

$$\mathfrak{G}\{f\} \leqq \mathfrak{G}\{w\} e^{\frac{\varepsilon}{\varrho}}.$$

On the other hand,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |z^n + \ldots|^p w(t)\, dt =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} |z^n + \ldots|^p f(t) \left\{ \frac{w(t)}{f(t)} - 1 \right\} dt + \frac{1}{2\pi} \int_{-\pi}^{\pi} |z^n + \ldots|^p f(t)\, dt.$$

Therefore, by virtue of Schwarz's inequality, we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |z^n + \ldots|^p w(t) \leqq$$

$$\leqq \frac{1}{2\pi} \sqrt{\int_{-\pi}^{\pi} |z^n + \ldots|^{2p} f^2(t)\, dt} \sqrt{\int_{-\pi}^{\pi} \left| \frac{w(t)}{f(t)} - 1 \right|^2 dt} +$$

$$+ \frac{1}{\sqrt{2\pi}} \sqrt{\int_{-\pi}^{\pi} |z^n + \ldots|^{2p} f^2(t)\, dt} \leqq$$

$$\leqq \frac{1}{\sqrt{2\pi}} \sqrt{\int_{-\pi}^{\pi} |z^n + \ldots|^{2p} f^2(t)\, dt} \cdot \left(1 + \frac{\varepsilon}{\varrho}\right)$$

and hence

$$\mu_n(w; p) \leqq \sqrt{\mu_n(f^2; 2p)} \left(1 + \frac{\varepsilon}{\varrho}\right).$$

Consequently,

$$\mu(w; p) \leqq \left(1 + \frac{\varepsilon}{\varrho}\right) \sqrt{\mu(f^2; 2p)} = \left(1 + \frac{\varepsilon}{\varrho}\right) \sqrt{\mathfrak{G}\{f^2\}} =$$

$$= \left(1 + \frac{\varepsilon}{\varrho}\right) \mathfrak{G}\{f\} \leqq \left(1 + \frac{\varepsilon}{\varrho}\right) e^{\frac{\varepsilon}{\varrho}} \mathfrak{G}\{w\}.$$

Since ε was arbitrarily chosen, the inequality (1) is established.

We now remove the restriction $w(t) \geq \varrho > 0$. Corresponding to a preassigned $\varepsilon > 0$, we obtain in every case

$$\mu(w + \varepsilon; p) \leqq \mathfrak{G}\{w + \varepsilon\}$$

and further

$$\mu(w; p) \leqq \mathfrak{G}\{w + \varepsilon\}.$$

Thereafter we have only to pass to the limit as $\varepsilon \to 0$; this gives no difficulty and will be omitted.

Thus the investigation of the case $\mathfrak{G}\{w\} > 0$ is complete.

Let us now assume $\mathfrak{G}\{w\} = 0$, i.e.

$$\int_{-\pi}^{\pi} \ln w(t)\, dt = -\infty.$$

For any $\varepsilon > 0$ we have

$$\mu(w + \varepsilon; p) = \mathfrak{G}\{w + \varepsilon\},$$

and it suffices to show that

$$\lim_{\varepsilon \to 0} \int_{-\pi}^{\pi} \ln [w(t) + \varepsilon]\, dt = -\infty .$$

For the proof, let us set

$$\int_{-\pi}^{\pi} \ln [w(t) + \varepsilon]\, dt = \int_{-\pi}^{\pi} \ln^+ [w(t) + \varepsilon]\, dt - \int_{-\pi}^{\pi} \ln^+ \frac{1}{w(t) + \varepsilon}\, dt = I_1(\varepsilon) - I_2(\varepsilon),$$

with

$$\ln^+ a = \begin{cases} \ln a, & \text{for } a > 1, \\ 0, & \text{for } 0 \leqq a \leqq 1. \end{cases}$$

Inasmuch as

$$0 \leqq I_1(\varepsilon) \leqq \int_{-\pi}^{\pi} [w(t) + \varepsilon]\, dt = \int_{-\pi}^{\pi} w(t)\, dt + 2\pi\varepsilon$$

it is necessary to prove that

$$\lim_{\varepsilon \to 0} I_2(\varepsilon) = \infty .$$

If we assume that

$$\lim_{\varepsilon \to 0} I_2(\varepsilon) = N < \infty$$

we find by Fatou's lemma [5] that

$$\int_{-\pi}^{\pi} \ln^+ \frac{1}{w(t)}\, dt \leqq N;$$

this, however, contradicts our hypothesis.

11. Let $\sigma(t)$ $(-\pi \leqq t \leqq \pi)$ be a non-decreasing function of bounded variation. Let $L^p(d\sigma;\ -\pi, \pi)$, with $p \geqq 1$, denote the totality of functions $f(t)$ such that

$$\int_{-\pi}^{\pi} |f(t)|^p\, d\sigma(t) < \infty\ .$$

$L^p(d\sigma;\ -\pi, \pi)$ will be a linear normalized space if we set

$$||f|| = \left\{ \int_{-\pi}^{\pi} |f(t)|^p\, d\sigma(t) \right\}^{\frac{1}{p}}$$

It is required to prove that the relation

$$\int_{-\pi}^{\pi} |\ln \sigma'(t)|\, dt = \infty$$

is necessary and sufficient for the closure of the sequence of functions

$$1,\quad e^{it},\quad e^{2it},\quad e^{3it},\ \ldots \tag{1}$$

in $L^p(d\sigma;\ -\pi, \pi)$. Here $\sigma'(t)$ is the derivative of the absolutely continuous part of the function $\sigma(t)$.

Kolmogoroff [90] was the first to prove the theorem for the case $p = 2$, and then Krein [91] proved it for the general case.

---

[5] Fatou's lemma asserts: if $g_k(x) \geqq 0$ $(k = 1, 2, \ldots)$ and if $g_k(x) \to g(x)$ almost everywhere in (a, b), it follows from

$$\int_a^b g_k(x)\, dx \leqq A \qquad (k = 1, 2, \ldots),$$

that

$$\int_a^b g(x)\, dx \leqq A$$

and hence, in particular, that $g(x)$ is integrable in (a, b).

Translator's Note: For a proof of the lemma see I. Natanson: *Theory of Functions of a Real Variable* (English by Boron and Hewitt), New York, 1955.

For the proof, we assume first of all that $\sigma(t)$ is an absolutely continuous function, that is we consider the space $L^p(wdt; -\pi, \pi)$, which will be denoted briefly by the symbol $L^p_w(-\pi, \pi)$. If

$$\int_{-\pi}^{\pi} |\ln w(t)|\, dt < \infty,$$

then it follows from Szegö's theorem that

$$\lim_{n\to\infty} \min_{A_k} \frac{1}{2\pi} \int_{-\pi}^{\pi} |e^{-it} - A_1 - A_2 e^{it} - \ldots - A_n e^{i(n-1)t}|^p\, w(t)\, dt =$$

$$= \lim_{n\to\infty} \min_{A_k} \frac{1}{2\pi} \int_{-\pi}^{\pi} |z^n - \bar{A}_1 z^{n-1} - \bar{A}_2 z^{n-2} - \ldots - \bar{A}_n|^p\, w(t)\, dt = \mathfrak{G}\{w\} > 0.$$

Hence the function $e^{-it} \in L^p_w(-\pi, \pi)$ cannot be approximated to the desired degree of accuracy by means of a linear combination of the functions (1). This proves the necessity of the condition.

We assume now that

$$\int_{-\pi}^{\pi} |\ln w(t)|\, dt = \infty\ .$$

In this case, by virtue of Szegö's theorem any one of the functions e^{-int} $(n = 1, 2, \ldots)$ can be approximated as closely as we please by means of a linear combination of the functions (1), and it remains only to remember that the set of functions

$$1,\ e^{\pm it},\ e^{\pm 2it}, \ldots$$

is closed in $L^p_w(-\pi, \pi)$.

We now prove that the criterion for the closure is still valid if $\sigma(t)$ is an arbitrary non-decreasing function of bounded variation and $w(t)$ denotes the derivative of its absolutely continuous part

$$d\sigma(t) = w(t)\, dt + d\sigma_{\text{disc}}(t) + d\sigma_{\text{sing}}(t)\ .$$

The necessity of the condition is obvious, since

$$\int_{-\pi}^{\pi} |f(t)|^p\, d\sigma(t) \geqq \int_{-\pi}^{\pi} |f(t)|^p\, w(t)\, dt.$$

In order to prove the sufficiency, let us assume that the system (1) is not closed in $L^p(d\sigma; -\pi, \pi)$. Then if, to be definite, we suppose $p > 1$, we obtain a function

$$g(t) \in L^q(d\sigma; -\pi, \pi) \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right),$$

such that

$$\int_{-\pi}^{\pi} g(t)\, e^{ikt}\, d\sigma(t) = 0 \qquad (k = 0, 1, 2, \ldots)$$

or

$$\int_{-\pi}^{\pi} e^{ikt}\, d\sigma^*(t) = 0 \qquad (k = 0, 1, 2, \ldots) \tag{2}$$

with

$$\sigma^*(t) = \int_{-\pi}^{t} g(\tau)\, d\sigma(\tau).$$

At this point we make use of a theorem due to F. Riesz [92], to the effect that the relations (2) can hold if and only if $\sigma^*(t)$ is an absolutely continuous function, i.e.

$$\sigma^*(t) = \int_{-\pi}^{\pi} g(t)\, w(t)\, dt.$$

Consequently the equations (2) have the form

$$\int_{-\pi}^{\pi} e^{ikt}\, g(t)\, w(t)\, dt = 0 \qquad (k = 0, 1, 2, \ldots),$$

whence it follows that the sequence (1) is not closed in $L_w^p(-\pi, \pi)$ either. This proves the theorem.

12. (Continuation). It is required to prove the set of functions

$$e^{iax} \qquad (a \geqq 0)$$

is closed in $L^p(d\sigma; -\infty, \infty)$ if and only if

$$\int_{-\infty}^{\infty} \frac{|\ln \sigma'(x)|}{1 + x^2}\, dx = \infty$$

(here $\sigma(x)$ is a non-decreasing function of bounded variation and $p \geq 1$).

For the case $p = 2$ the theorem has been proved by Krein [93].
If $f(x) \in L^p(d\sigma; -\infty, \infty)$, then

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \left| f(x) - \sum_{k=0}^{n} \frac{A_k}{(x+i)^k} \right|^p d\sigma(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \left| f(x) - \sum_{k=0}^{n} B_k \left(\frac{x-i}{x+i} \right)^k \right|^p d\sigma(x) =$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| f\left(tg\, \frac{t}{2} \right) - \sum_{k=0}^{n} C_k e^{ikt} \right|^p d\sigma\left(tg\, \frac{t}{2} \right).$$

Therefore the sequence of functions

$$1, \frac{1}{x+i}, \frac{1}{(x+i)^2}, \dots \tag{3}$$

is closed in $L^p(d\sigma; -\infty, \infty)$ if and only if the sequence

$$1, e^{it}, e^{2it}, e^{3it}, \dots$$

is closed in $L^p(d\sigma_1; -\pi, \pi)$ with $\sigma_1(t) = \sigma\left(\operatorname{tg}\frac{t}{2}\right)$.

By reason of the result in Section **11**, the condition

$$\int_{-\infty}^{\infty} \frac{|\ln \sigma'(x)|}{1+x^2}\, dx = \infty \tag{4}$$

is equivalent to the condition

$$\int_{-\pi}^{\pi} |\ln \sigma_1'(t)|\, dt = \infty .$$

Consequently this condition is necessary and sufficient for the closure of the sequence (3), in $L^p(d\sigma; -\infty, \infty)$.

Let us consider the development

$$\frac{i^n}{(x+i)^n} = \frac{1}{(n-1)!} \int_0^{\infty} u^{n-1} e^{iu(x+i)}\, du \qquad (n = 1, 2, \dots).$$

Corresponding to any preassigned $\varepsilon > 0$, there exists an $N > 0$ so that

$$\left| \frac{i^n}{(x+i)^n} - \frac{1}{(n-1)!} \int_0^{N} u^{n-1} e^{iu(x+i)}\, du \right| < \frac{\varepsilon}{2\left\{ \int_{-\infty}^{\infty} d\sigma(x) \right\}^{\frac{1}{p}}} \qquad (-\infty < x < \infty) .$$

Therefore

$$\int_{-\infty}^{\infty} \left| \frac{i^n}{(x+i)^n} - \frac{1}{(n-1)!} \int_0^{N} u^{n-1} e^{iu(x+i)}\, du \right|^p d\sigma(x) < \left(\frac{\varepsilon}{2}\right)^p . \tag{5}$$

If we set

$$\varphi(x) = \frac{1}{(n-1)!} \int_0^{N} u^{n-1} e^{iu(x+i)}\, du ,$$

$$\varphi_m(x) = \frac{1}{(n-1)!} \frac{N}{m} \sum_{k=0}^{m-1} \left(\frac{kN}{m}\right)^{n-1} e^{-\frac{kN}{m}} e^{i\frac{kN}{m}x} ,$$

we find

$$|\varphi(x) - \varphi_m(x)| < 1 + e^N \qquad (-\infty < x < \infty;\ m = 1, 2, \ldots),$$

and consequently

$$\int_{-\infty}^{\infty} |\varphi(x) - \varphi_m(x)|^p \, d\sigma(x) < K.$$

Here K is independent of m. But since, for all values of x,

$$\lim_{m\to\infty} \varphi_m(x) = \varphi(x)$$

it follows that

$$\lim_{m\to\infty} \int_{-\infty}^{\infty} |\varphi(x) - \varphi_m(x)|^p \, d\sigma(x) = 0.$$

Therefore we can find an M such that

$$\int_{-\infty}^{\infty} |\varphi(x) - \varphi_M(x)|^p \, d\sigma(x) < \left(\frac{\varepsilon}{2}\right)^p.$$

Hence, by considering (5) we get

$$\int_{-\infty}^{\infty} \left| \frac{i^n}{(x+i)^n} - \varphi_M(x) \right|^p d\sigma(x) < \varepsilon^p.$$

But since $\varphi_M(x)$ is a linear combination of the functions

$$e^{iax} \qquad (a \geqq 0) \tag{6}$$

we can approximate every one of the "polynomials"

$$\sum_{k=0}^{n} A_k \frac{1}{(x+i)^k}$$

in $L^p(d\sigma; -\infty, \infty)$ as closely as we please through a linear combination of the functions (6). Hence the closure of the sequence (3) in $L^p(d\sigma; -\infty, \infty)$ implies the closure of the set (6), that is, the condition (4) is sufficient for the closure of the set (6).

It remains only to prove the necessity of this condition.

For definiteness let us assume $p > 1$. Suppose

$$\int_{-\infty}^{\infty} \frac{|\ln \sigma'(x)|}{1+x^2} \, dx < \infty.$$

Then there exists a function $g(x) \in L^p(d\sigma; -\infty, \infty)$, with $\frac{1}{p} + \frac{1}{q} = 1$, such that

$$\int_{-\infty}^{\infty} \frac{g(x)}{(x+i)^n}\, d\sigma(x) = 0 \qquad (n = 0, 1, 2, \ldots) \,. \tag{7}$$

From (7) we obtain

$$\int_{-\infty}^{\infty} \frac{g(x)}{x+z}\, d\sigma(x) = 0$$

for all values of z in the upper half-plane. This relation can be brought into the form

$$\int_{-\infty}^{\infty} g(x)\, d\sigma(x) \int_{0}^{\infty} e^{iu(x+z)}\, du = 0$$

or

$$\int_{0}^{\infty} e^{iuz}\, du \int_{-\infty}^{\infty} e^{iux} g(x)\, d\sigma(x) = 0 \,.$$

By reason of the formula derived for Fourier integrals, we have almost everywhere on the semi-axis $a \geq 0$:

$$\int_{-\infty}^{\infty} e^{iax} g(x)\, d\sigma(x) = 0;$$

but since the left-hand side is continuous, this equation holds for every $a \geq 0$. Whence it follows that the set of functions $e^{iax} (a \geq 0)$ is not closed in $L^p(d\sigma; -\infty, \infty)$.

13. Let $s(x)$ be a function which is positive and continuous in the interval $[-1, 1]$ and let

$$\mathscr{L}_n[s(x)] = \min_{A_k} \max_{-1 \leq x \leq 1} |s(x)\,(x^n + A_1 x^{n-1} + \ldots + A_n)| \,.$$

Then:

$$\lim_{n \to \infty} \frac{\mathscr{L}_n[s(x)]}{\frac{1}{2^{n-1}}} = e^{\frac{1}{\pi} \int_{-1}^{1} \frac{\ln s(x)}{\sqrt{1-x^2}}\, dx}$$

This continuation is connected with the theorem in Section **10** and is due to Bernstein.

C. Further Examples of Closed Sequences of Functions

14. It is required to prove that each of the sequences of functions

$$\{\operatorname{sign}\sin kt\}_{k=1}^{\infty}, \qquad \{\operatorname{sign}\cos kt\}_{k=0}^{\infty}$$

is closed in $L^2(0, \pi)$.

We give the proof for the first sequence. It suffices to show that from $f(t) \in L^2(0, \pi)$ and the equations

$$\int_0^{\pi} f(t) \operatorname{sign}\sin kt\, dt = 0 \qquad (k = 1, 2, 3, \ldots) \tag{1}$$

we get the equations $c_k = 0$ $(k = 1, 2, \ldots)$, where

$$c_k = \frac{2}{\pi}\int_0^{\pi} f(t) \sin kt\, dt.$$

Since $f(t) \in L^2(0, \pi)$, we have

$$\sum_{k=1}^{\infty} |c_k|^2 < \infty. \tag{2}$$

On the other hand, from the development

$$\operatorname{sign}\sin kt = \frac{4}{\pi}\sum_{r=0}^{\infty} \frac{\sin (2r+1)kt}{2r+1} \qquad (k = 1, 2, 3, \ldots)$$

and the relations (1) we find the equations

$$\sum_{r=0}^{\infty} \frac{c_{k(2r+1)}}{2r+1} = 0 \qquad (k = 1, 2, 3, \ldots). \tag{3}$$

We must therefore prove that the relations (2) and (3) are possible only when $c_k = 0$ $(k = 1, 2, \ldots)$.

Let us assume the contrary, that is, we assume $c_k \neq 0$. Out of the totality of all the equations (3), let us consider those for which $k = n(2s+1)$, $(s = 0, 1, 2, \ldots)$; we will represent the latter group by

$$\sum_{r=0}^{\infty} \frac{c_{n(2s+1)(2r+1)}}{(2s+1)(2r+1)} = 0 \quad .(s = 0, 1, 2, \ldots) \; . \tag{4}$$

Let $\tau(m)$ denote the number of divisors of the number m.

As we know [94] we then have for any $\varepsilon > 0$:

$$\lim_{m\to\infty} \frac{\tau(m)}{m^{\varepsilon}} = 0.$$

Hence it follows from (2) that the series

$$\sum_{m=1}^{\infty} \frac{\tau(m)\, c_m}{m}$$

is absolutely convergent. Then the double series

$$\sum_{i,\, k=1}^{\infty} \frac{c_{ik}}{ik}$$

also is absolutely convergent.

Accordingly, the same properties belong to the series

$$\sum_{r,\, s=0}^{\infty} \mu(2s+1)\, \frac{c_{n(2s+1)(2r+1)}}{(2s+1)(2r+1)},$$

where $\mu(a)$ is the Möbius function; that is, $\mu(1) = 1$, $\mu(a) = 0$ if a has quadratic factors other than unity, and $\mu(a) = (-1)^{\nu}$ if a has no quadratic factors (excluding unity); here ν is the number of prime factors of a (unity not being included among the prime factors).

From the foregoing it follows by (4) that

$$0 = \sum_{s=0}^{\infty} \mu(2s+1) \sum_{r=0}^{\infty} \frac{c_{n(2s+1)(2r+1)}}{(2s+1)(2r+1)} = \sum_{m=0}^{\infty} \left\{ \sum_{d \mid 2m+1} \mu(d) \right\} \frac{c_{n(2m+1)}}{2m+1};$$

here d runs through the set of all factors of the numbers $2m+1$. But since

$$\sum_{d \mid 2m+1} \mu(d) = \begin{cases} 1 & (m = 0), \\ 0 & (m > 0) \end{cases}$$

we have

$$c_n = 0,$$

and the proposition is established.

15. Let

$$v_k(t) = \operatorname{sign} \sin kt, \qquad (k = 1, 2, 3, \ldots),$$

$$u_k(t) = \frac{1}{2} \sum_{d_1 \mid k} \frac{\mu(d_1)}{d_1} \sin \frac{kt}{d_1}$$

where d_1 runs through the set of all odd factors of k, and $\mu(\alpha)$ is the Möbius function (as defined in Section **14**). It is required to prove the equation

$$\int_0^{\pi} u_i(t)\, v_k(t)\, dt = \begin{cases} 1 & (i = k), \\ 0 & (i \neq k). \end{cases}$$

We have here a biorthogonal system which was first studied by Tchebysheff [35].

16. Let

$$v_0(t) = 1, \qquad v_k(t) = \operatorname{sign} \cos kt, \qquad (k = 1, 2, 3, \ldots),$$

$$u_0(t) = \frac{1}{\pi}, \qquad u_k(t) = \frac{1}{2} \sum_{d_1 \mid k} \frac{(-1)^h}{d_1} \cos \frac{kt}{d_1}$$

where d_1 runs through the set of all odd factors of k, quadratic factors being excluded (unity does not count as a quadratic factor), and h is the number of factors of the form $4m + 1$ which divide d_1 without remainder. It is required to prove the following equation:

$$\int_0^{\pi} u_i(t)\, v_k(t)\, dt = \begin{cases} 1 & (i = k), \\ 0 & (i \neq k). \end{cases}$$

This biorthogonal system is due to A. A. Markoff [36].

17. It is required to prove that every sequence consisting of an infinite set of Bernoulli polynomials $B_n(x)$ of even degree, including $B_0(x) = 1$, is closed in the space $C\left(0, \frac{1}{2}\right)$.

This theorem, due to Müntz, is easily proved with the aid of the development in Section 55 and the following remark: Suppose that

$$D_\varrho(x) = \sum_{\nu=1}^{\infty} \frac{\sin 2\pi\nu x}{\nu^\varrho} \qquad (\varrho > 1)$$

and that $f(x) \in L^2\left(0, \frac{1}{2}\right)$; then the function

$$F(\varrho) = \int_0^{\frac{1}{2}} f(x)\, D_\varrho(x)\, dx = \sum_{\nu=1}^{\infty} \frac{c_\nu}{\nu^\varrho},$$

with

$$c_\nu = \int_0^{\frac{1}{2}} f(x) \sin 2\pi\nu x\, dx,$$

is certainly different from zero for all sufficiently large values of ϱ, provided not all the quantities c_ν are equal to zero.

18. It is required to prove [95] that every sequence containing an infinite set of Bernoulli polynomials of even degree as well as an infinite set of Bernoulli polynomials of odd degree, inclusive of $B_0(x)$ and $B_1(x)$, is closed in the space $C(0, 1)$.

D. The Carathéodory-Fejér Problem and Similar Problems

19. Let $K_p (p = 0, 1, 2, \ldots)$ be the set of all analytic functions $f(z)$ which are regular in the circle $|z| \leq 1$ with the possible exception of certain poles lying in the region $|z| < 1$, the number of these poles, counted correctly (i.e. with due regard for multiplicity), being not greater than p.

In addition let

$$F(z) = \frac{c_0}{z^{n+1}} + \frac{c_1}{z^n} + \ldots + \frac{c_n}{z} \qquad (c_0 \neq 0).$$

Then the expression

$$\inf_{f \in K_p} \max_{z=1} |F(z) - f(z)| = \mu_p[F] \qquad (p = 0, 1, \ldots, n)$$

cannot be equal to zero.

Suppose there exists a rational function $\varphi(x)$ that has exactly p poles in the region $|z| < 1$ and is regular at the point $z = 0$, so that

$$z^{n+1}\{F(z) - \varphi(z)\} = \lambda \frac{a_0 + a_1 z + \ldots + a_m z^m}{\bar{a}_m + \bar{a}_{m-1} z + \ldots + \bar{a}_0 z^m}$$

with $m \leq n$ and $\lambda > 0$.

Under these conditions, we are required to prove [96]:

$$\mu_p[F] = \mu_{p+1}[F] = \ldots = \mu_{p+n-m}[F] = \lambda .$$

Since

$$\max_{|z|=1} |F(z) - \varphi(z)| = \lambda$$

and since $\zeta(z)$ belongs to the class K_r for $r \geq p$, we have

$$\mu_r[F] \leq \lambda \qquad (r \geq p).$$

Hence it suffices to prove that

(1) $$\max_{|z|=1} |F(z) - f(z)| \geq \lambda$$

for every function $f(z) \in K_{p+n-m}$. Suppose, then, that $f(z) \in K_{p+n-m}$. Let us set

$$R(z) = \lambda \frac{a_0 + a_1 z + \ldots + a_m z^m}{\bar{a}_m + \bar{a}_{m-1} z + \ldots + \bar{a}_0 z^m}$$

and

$$z^{n+1}\{F(z) - f(z)\} = z^{n+1}\{F(z) - \varphi(z)\} + z^{n+1}\{\varphi(z) - f(z)\} = R(z) + \psi(z).$$

In addition, let σ denote the number of poles of $\psi(z)$ in the region $|z| < 1$, and let τ denote the number of zeros of $\psi(z)$ in this region.

Then we have

$$0 \leq \sigma \leq 2p + n - m - j, \qquad \tau \geq n + 1 - j,$$

where j $(0 \leq j \leq p + n - m)$ is the order of multiplicity of a possible pole of the function $f(z)$ at $z = 0$.

Since the function $R(z)$ has exactly p poles in the region $|z| < 1$, and since $R(z)$ is a constant on the circumference of the circle $|z| = 1$, it follows that the number of zeros of the function $R(z)$ in the region $|z| < 1$ is not greater than $m - p$.

Let us assume first that

$$\min_{|z|=1} |\psi(z)| = \delta > 0 .$$

If z describes the circle $|z| = 1$ in a positive sense, then

$$\arg R(z) \quad \text{and} \quad \arg \psi(z)$$

increase by the amounts $2\pi\alpha$ and $2\pi\beta$, respectively, where, by the foregoing,

$$\alpha \leqq m - p - p = m - 2p, \qquad \beta = \tau - \sigma \geqq 1 + m - 2p .$$

Because of the relation

$$\beta - \alpha \geqq 1$$

we can find on the circumference of the circle $|z| = 1$ at least one point ζ at which

$$\arg R(\zeta) = \arg \psi(\zeta) .$$

Consequently,

$$|F(\zeta) - f(\zeta)| = |R(\zeta) + \psi(\zeta)| = |R(\zeta)| + |\psi(\zeta)| \geqq \lambda + \delta,$$

and thus the inequality (1) is established for this case.

Now let us assume that along the circle $|z| = 1$ there exist points at which $\psi(z) = 0$. Since, by hypothesis, the function in question is regular along the circle $|z| = 1$, there exists a ϱ $(0 < \varrho < 1)$ such that $\psi(z)$, $R(z)$ are regular in the region $\varrho < |z| < 1$ and different from zero in this region. If we preassign an r which satisfies the condition $\varrho < r < 1$ and describes the circle $|z| = r$, we find, as before, at least one point ζ_r at which we have:

$$|F(\zeta_r) - f(\zeta_r)| > |R(\zeta_r)| .$$

If we cause r to approach unity, we come to a certain point ζ $(|\zeta| = 1)$ at which

$$|F(\zeta) - f(\zeta)| \geqq |R(\zeta)| = \lambda$$

and hence the inequality (1) is valid also in this case.

20. (Continuation.) Let

$$F(z) = \frac{c_0}{z^{n+1}} + \frac{c_1}{z^n} + \ldots + \frac{c_n}{z} \qquad (c_0 \neq 0) .$$

Let us set

$$D_n(\lambda) = \begin{vmatrix} \lambda & 0 & \dots & 0 & c_0 & c_1 & \dots & c_n \\ 0 & \lambda & \dots & 0 & 0 & c_0 & \dots & c_{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda & 0 & 0 & \dots & c_0 \\ \bar{c}_0 & 0 & \dots & 0 & \lambda & 0 & \dots & 0 \\ \bar{c}_1 & \bar{c}_0 & \dots & 0 & 0 & \lambda & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{c}_n & \bar{c}_{n-1} & \dots & \bar{c}_0 & 0 & 0 & \dots & \lambda \end{vmatrix}.$$

All the roots of this polynomial are real, pairwise symmetrical and different from zero. Suppose the positive roots

$$\lambda_0 > \lambda_1 > \dots > \lambda_q$$

of the polynomial $D_n(x)$ have the multiplicities

$$\nu_0, \nu_1, \dots, \nu_q \qquad (\nu_0 + \nu_1 + \dots + \nu_q = n + 1).$$

It is required to prove the equations

$$\mu_p[F] = \lambda_j \qquad (p = 0, 1, \dots, n)$$

where the subscript j is determined by the inequalities

$$\nu_0 + \nu_1 + \dots + \nu_{j-1} \leqq p < \nu_0 + \nu_1 + \dots + \nu_j$$

and $j = 0$ when $p \leq \nu_0 - 1$.

Section **19** naturally suggests seeking fractions of the form

$$R(z) = \lambda \frac{a_0 + a_1 z + \dots + a_n z^n}{\bar{a}_n + \bar{a}_{n-1} z + \dots + \bar{a}_0 z^n} \qquad (\lambda > 0) \tag{2}$$

that allow the development

$$R(z) = c_0 + c_1 z + c_2 z^2 + \dots + c_n z^n + \gamma_{n+1} z^{n+1} + \dots \tag{2'}$$

in a neighborhood of the point $z = 0$. This leads to the system of equations

$$\begin{cases} \lambda a_0 = c_0 \bar{a}_n, \\ \lambda a_1 = c_1 \bar{a}_n + c_0 \bar{a}_{n-1}, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \lambda a_n = c_n \bar{a}_n + h_{n-1} \bar{a}_{n-1} + \dots + c_0 \bar{a}_0. \end{cases} \tag{3}$$

Whence it follows that λ must be a root of the polynomial $D_n(\lambda)$.

Let $\lambda^* > 0$ be a root of the polynomial $D_n(\lambda)$. This root corresponds to a certain totality of solutions of the system (3) and consequently also to a certain set of rational fractions of the form (2). But do the above fractions allow a development of the form (2′) in the neighborhood of the point $z = 0$?

We will show that the answer to the question just raised is in the affirmative, provided certain conditions are fulfilled. We will prove namely: If λ^* is a root of the polynomial $D_n(\lambda)$ and if $D_{n-1}(\lambda^*) \neq 0$, there exists a fraction

$$R^*(z) = \lambda^* \frac{a_0 + a_1 z + \ldots + a_n z^n}{\bar{a}_n + \bar{a}_{n-1} z + \ldots + \bar{a}_0 z^n}, \tag{4}$$

which allows a development of the form (2′) and, besides, is irreducible.

In fact, by reason of the condition $D_{n-1}(\lambda^*) \neq 0$, the matrix corresponding to the system (3) is of rank $2n + 1$ when $\lambda = \lambda^*$. Hence the totality of solutions of the system (3) for $\lambda = \lambda^*$ is given by the formulas

$$\sigma_k = \sigma c_k^* \qquad (k = 0, 1, \ldots, n)$$

where σ is an arbitrary real quantity and a_k^* $(k = 1, 2, \ldots, n)$ are the algebraic complements of the first row of the determinant $D_n(\lambda^*)$. The important fact is that

$$a_n^* = \lambda^* D_{n-1}(\lambda^*) \neq 0$$

and consequently

$$a_0^* \neq 0 .$$

Therefore the fraction $R^*(z)$, corresponding to a solution of the system (3) for $\lambda = \lambda^*$, has the form (4), where $a_0 \neq 0$, $a_n \neq 0$. If we develop this fraction according to powers of z, we find that the first $n + 1$ coefficients of this development, which are related to the numbers $\lambda^*, a_0, \ldots, a_n$ by equations, are identical with the equations (3). Hence the development of the fraction $R^*(z)$ has the form (2′).

If the fraction $R^*(z)$ were reducible, we could write it in the form

$$R^*(z) = \lambda^* \frac{b_0 + b_1 z + \ldots + b_{n-1} z^{n-1}}{\bar{b}_{n-1} + \bar{b}_{n-2} z + \ldots + \bar{b}_0 z^{n-1}}$$

where $b_0 \neq 0$, $b_{n-1} \neq 0$. By comparing this fraction with (2′), we obtain

$$\begin{cases} \lambda^* b_0 = c_0 \bar{b}_{n-1}, \\ \lambda^* b_1 = c_1 \bar{b}_{n-1} + c_0 \bar{b}_{n-2}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \lambda^* b_{n-1} = c_{n-1} \bar{b}_{n-1} + c_{n-2} \bar{b}_{n-2} + \ldots + c_0 \bar{b}_0, \\ 0 = c_n \bar{b}_{n-1} + c_{n-1} \bar{b}_{n-2} + \ldots + c_1 \bar{b}_0, \end{cases}$$

whence it would follow that $D_{n-1}(\lambda^*) = 0$.

In order to arrive at the proof of our theorem, let us assume first that all the roots of the polynomial $D_n(\lambda)$ are simple and that none of them is a root of the polynomial $D_{n-1}(\lambda)$. Let us denote the positive roots of the polynomial $D_n(\lambda)$ by

$$\lambda_0 > \lambda_1 > \ldots > \lambda_n .$$

From what we have already proved, to every root λ_k there corresponds an irreducible rational fraction $R_k(z)$ of the form (2), allowing the development (2′). The results in Section **19** show that the number of poles of the fraction $R_{k+1}(z)$ in the region $|z| < 1$, must be greater than the number of poles of the fraction $R_k(z)$. Therefore $R_k(z)$ has exactly k poles in the region $|z| < 1$, and thus

$$\lambda_k = \mu_k[F] \qquad (k = 0, 1, \ldots, n) .$$

We come now to the consideration of the excluded cases. It is easy to see that we can always construct a continuous function $c_i(\varepsilon)$ $(i = 1, 2, \ldots, n)$ of the parameters $\varepsilon > 0$, such that,

in the first place, $c_i(0) = c_i$ and, secondly, the polynomial $D_n(\lambda; \varepsilon)$ for $\varepsilon > 0$ has only simple roots, none of them a root of the polynomial $D_{n-1}(\lambda; \varepsilon)$. If $\varepsilon \to 0$, the positive roots

$$\lambda_0(\varepsilon) > \lambda_1(\varepsilon) > \ldots > \lambda_n(\varepsilon)$$

of the polynomial $D_n(\lambda; \varepsilon)$ fulfill the relations

$$\lambda_p(\varepsilon) \to \lambda_j,$$

where

$$(5) \qquad \nu_0 + \nu_1 + \ldots + \nu_{j-1} \leqq p < \nu_0 + \nu_1 + \ldots \nu_j$$

with $j = 0$ for $p \leq \nu_0 - 1$.

Let us set

$$F_\varepsilon(z) = \frac{c_0 + c_1(\varepsilon)\, z + \ldots + c_n(\varepsilon)\, z^n}{z^{n+1}} = F(z) + \frac{P_{n-1}(z; \varepsilon)}{z}.$$

Then we have

$$-\max_{|z|=1} |P_{n-1}(z; \varepsilon)| \leqq \mu_p[F_\varepsilon] - \mu_p[F] \leqq \max_{|z|=1} |P_{n-1}(z; \varepsilon)|;$$

therefore

$$\lambda_p(\varepsilon) - \max_{|z|=1} |P_{n-1}(z;\varepsilon)| \leqq \mu_p[F] \leqq \lambda_p(\varepsilon) + \max_{|z|=1} |P_{n-1}(z; \varepsilon)|,$$

and consequently we find

$$\mu_p[F] = \lambda_j,$$

where j is defined by (5). This completes the proof of the proposition.

It is worthwhile noting that $D_n(\lambda)$ is the discriminant of the Hermitian form

$$H_n(\lambda) = \lambda^2 \sum_{k=1}^{n+1} |x_k|^2 - \sum_{k=1}^{n+1} |c_0 x_k + c_1 x_{k+1} + \ldots + c_{n+1-k} x_{n+1}|^2 .$$

21. With the results in Section **20** as a starting point, let us solve the following problem proposed by Carathéodory and Fejér [97]:

Given the numbers

$$c_0, c_1, c_2, \ldots, c_n.$$

Of all the functions

$$g(z) = c_0 + c_1 z + \ldots + c_n z^n + \gamma_{n+1} z^{n+1} + \ldots$$

that are regular in the region $|z| < 1$, it is required to find the particular one for which

$$\sup_{|z|<1} |g(z)|$$

is a minimum.

That is, we are to determine

$$\mu = \min \sup_{z=1} |g(z)| .$$

Let

$$F(z) = \frac{c_0}{z^{n+1}} + \frac{c_1}{z^n} + \ldots + \frac{c_n}{z}.$$

We know that

$$\inf_{f \in K_n} \max_{|z|=1} |F(z) - f(z)| = \mu_0 [F] = \lambda_0$$

is the largest root of the polynomial $D_n(\lambda)$. Let us prove that, for this (largest) root λ_0 there exists a rational fraction

$$(6) \qquad R_0(z) = \lambda_0 \frac{a_0 + a_1 z + \ldots + a_n z^n}{\bar{a}_n + \bar{a}_{n-1} z + \ldots + \bar{a}_0 z^n} = c_0 + c_1 z + \ldots + c_n z^n + \gamma_{n+1} z^{n+1} + \ldots$$

We have already proved this by assuming that λ_0 is a simple root of $D_n(\lambda)$ with $D_{n-1}(\lambda_0) \neq 0$. If this hypothesis be abandoned, we can introduce a function $c_i(\varepsilon)$ of the parameter $\varepsilon > 0$, just as before. Corresponding to every $\varepsilon > 0$, there exists a rational fraction $R_0(z; \varepsilon)$ having the coefficients $a_k(\varepsilon)$ and the maximum absolute value $\lambda_0(\varepsilon)$.

We may assume that

$$\sum_{k=0}^{n} |a_k(\varepsilon)| = 1 .$$

Then there exists a sequence of values $\varepsilon_i \to 0$ for which we have:

$$a_n(\varepsilon_i) \to a_k \qquad (k = 0, 1, \ldots, n) .$$

After passing to the limit, we obtain a rational function $R_0(z)$ which assumes the constant value λ_0 along the circle $|z| = 1$, is regular within the circle $|z| \leq 1$ and has the development

$$c_0 + c_1 z + \ldots + c_n z^n + \gamma_{n+1} z^{n+1} + \ldots$$

Let $g(z)$ be any function which is regular in the region $|z| < 1$ and has the development

$$(7) \qquad g(z) = c_0 + c_1 z + \ldots + c_n z^n + \ldots$$

We are going to prove the relation

$$\sup_{|z|<1} |g(z)| \geqq \max_{|z|=1} |R_0(z)| = \lambda_0$$

and show that the equality sign holds only when $g(z) = R_0(z)$. Then it will follow that the Carathéodory-Fejér extremal function is a rational function $R_0(z)$ and further, that the quantity λ_0 is the largest root of the polynomial $D_n(\lambda)$.

Let us form the product

$$\varphi(z) = g(z)(\bar{a}_n + \bar{a}_{n-1} z + \ldots + \bar{a}_0 z^n).$$

This function is regular in the region $|z| < 1$ and, because of (6), (7), has the development

$$\varphi(z) = \lambda_0(a_0 + a_1 z + \ldots + a_n z^n + b_{n+1} z^{n+1} + \ldots).$$

For $0 \leq \varrho < 1$ we find

$$\frac{1}{2\pi}\int_0^{2\pi} |g(\varrho e^{i\Theta})(\bar{a}_n + \bar{a}_{n-1}\varrho e^{i\Theta} + \ldots + \bar{a}_0 \varrho^n e^{ni\Theta})|^2 d\Theta =$$

$$= \lambda_0^2 \sum_{k=0}^{n} |a_k|^2 \varrho^{2k} + \lambda_0^2 \sum_{k=n+1}^{\infty} |b_k|^2 \varrho^{2k}.$$

Hence it follows from

$$\sup_{|z|<1} |g(z)| = G,$$

that

$$\lambda_0^2 \sum_{k=0}^{n} |a_k|^2 \varrho^{2k} + \lambda_0^2 \sum_{k=n+1}^{\infty} |b_k|^2 \varrho^{2k} \leqq G^2 \sum_{k=0}^{n} |a_k|^2 \varrho^{2k}.$$

We see: Provided not all the quantities b_k are equal to zero, G is surely greater than λ_0. This proves the proposition.

22. It is required to prove the following theorem due to I. Schur [98]: In order that the power series

$$\sum_{k=0}^{\infty} c_k z^k$$

represent the function $g(x)$, which satisfies the inequality

$$|g(z)| < 1$$

in the circle $|z| < 1$, it is necessary and sufficient that for an arbitrary n the Hermitian form

$$H = \sum_{k=1}^{n+1} |x_k|^2 - \sum_{k=1}^{n+1} |c_0 x_k + c_1 x_{k+1} + \ldots + c_{n+1-k} x_{n+1}|^2$$

be non-negative.

The condition given in the statement of the problem means in substance that the largest eigenvalue of the form

$$\sum_{k=1}^{n+1} |c_0 x_k + c_1 x_{k+1} + \ldots + c_{n+1-k} x_{n+1}|^2,$$

that is the square of the largest root of the polynomial $D_n(\lambda)$ for all values of n, is not greater than unity.

If this condition is fulfilled, there exists for an arbitrary n a function

$$R_0(z;n) = c_0 + c_1 z + \ldots + c_n z^n + \gamma_{n+1}^{(n)} z^{n+1} + \gamma_{n+2}^{(n)} z^{n+2} + \ldots$$

whose numerical value in the circle $|z| \leq 1$ is not greater than unity.

By the well-known Vitali theorem we can select from the sequence

$$\{R_0(z;n)\}_{n=0}^{\infty}$$

a sub-sequence with the following properties: it is uniformly convergent in the interior of the region $|z| < 1$; its limit function obviously has the development

$$c_0 + c_1 z + c_2 z^2 + \ldots$$

and its value in the region $|z| < 1$ does not exceed unity.

The proof of the necessity of the condition is even simpler.

Indeed if the function

$$g(z) = c_0 + c_1 z + c_2 z^2 + \ldots$$

satisfies the inequality

$$|g(z)| < 1 \qquad (|z| < 1)\ ,$$

it follows from the results in Section **21** that the largest root of $D_n(\lambda)$ is not greater that unity.

23. With the results in Sections **19, 20** as a basis, let us prove the following theorem due to A. Cohn [99]: If the Hermitian form

$$H = \sum_{i=1}^{n} |\bar{a}_n z_i + \bar{a}_{n-1} z_{i+1} + \ldots + \bar{a}_n z_n|^2 - \sum_{i=1}^{n} |a_0 z_i + a_1 z_{i+1} + \ldots + a_{n-i} z_n|^2,$$

where $a_0 \neq 0$, $a_n \neq 0$, has the rank n and the signature $2m - n$, then the equation

$$a_0 + a_1 z + \ldots + a_n z^n = 0$$

has exactly m roots in the region $|z| < 1$ and exactly $n - m$ roots in the region $|z| > 1$.

For the proof, let us consider the system of equations

$$\begin{cases} a_0 = c_0 \bar{a}_n, \\ a_1 = c_0 \bar{a}_{n-1} + c_1 a_n, \\ \cdots\cdots\cdots \\ a_n = c_0 \bar{a}_0 + c_1 \bar{a}_1 + \ldots + c_n a_n \end{cases}$$

and determine the numbers c_i in terms of the a_i's.

By setting

$$x_k = \bar{a}_n z_k + \bar{a}_{n-1} z_{k+1} + \ldots + \bar{a}_k z_n + \bar{a}_{k-1} z_{n+1} \qquad (k = 1, 2, \ldots, n+1)$$

we can prove without difficulty that

$$\sum_{i=1}^{n} |\bar{a}_n z_i + \bar{a}_{n-1} z_{i+1} + \ldots + \bar{a}_i z_n|^2 - \sum_{i=1}^{n} |a_0 z_i + a_1 z_{i+1} + \ldots + a_{n-i} z_n|^2 =$$

$$= \sum_{k=1}^{n+1} |x_k|^2 - \sum_{k=1}^{n+1} |c_0 x_k + c_1 x_{k+1} + \ldots + c_{n+1-k} x_{n+1}|^2 = H_n(1).$$

Together with the new form and the quantities c_i $(i = 1, 2, \ldots, n)$, let us consider the function

$$R(z) = \frac{a_0 + a_1 z + \ldots + a_n z^n}{\bar{a}_n + \bar{a}_{n-1} z + \ldots + \bar{a}_0 z^n} = c_0 + c_1 z + \ldots + c_n z^n + \gamma_{n+1} z^{n+1} + \ldots .$$

The polynomial $D_n(\lambda)$ has the root $\lambda = 1$. The positive roots of this polynomial are, say,

$$\lambda_0 \geqq \lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n.$$

The squares of these numbers are eigenvalues of the form

$$\sum_{k=1}^{n+1} |c_0 x_k + c_1 x_{k+1} + \ldots + c_{n+1-k} x_{n+1}|^2.$$

Therefore, by virtue of our hypothesis, we have

$$\lambda_{n-m+1} < 1 = \lambda_{n-m} < \lambda_{n-m-1}.$$

Hence it follows from the results in Section **20** that

$$\inf_{f \in K_{n-m}} \max_{z=1} \left| \frac{c_0}{z^{n+1}} + \frac{c_1}{z^n} + \ldots + \frac{c_n}{z} - f(z) \right| \equiv \mu_{n-m}[F] = 1,$$

while

$$\mu_{n-m-1}[F] > 1, \quad \mu_{n-m+1}[F] < 1 .$$

Therefore, from the results in Section **19**, the fraction $R(z)$ is irreducible and has exactly $n - m$ poles in the region $|z| < 1$. Since the fraction $R(z)$ is irreducible, its denominator cannot vanish along the circle $|z| = 1$, and hence $R(z)$ has exactly m zeros in the region $|z| > 1$. The only thing still left is to observe that the zeros of the numerators and those of the denominators are central-symmetric with respect to the circle $|z| = 1$.

24. Given, the real numbers $a_0 \neq 0$, a, $\ldots$, a_n and an integer $N > n$. To prove:

$$\min_{p_i, q_i} \max_{-1 \leqq x \leqq 1} \left| a_0 x^N + \frac{a_1}{2} x^{N-1} + \ldots + \frac{a_n}{2^n} x^{N-n} - \right.$$

$$\left. - \frac{q_0 x^{N-1} + q_1 x^{N-2} + \ldots + q_{N-1}}{p_0 x^n + \ldots + p_n} \right| = \frac{|\lambda|}{2^{N-1}}$$

where λ is numerically the smallest root of the polynomial

$$\begin{vmatrix} c_n - \lambda & c_{n-1} & \dots & c_1 & c_0 \\ c_{n-1} & c_{n-2} - \lambda & \dots & c_0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_1 & c_0 & \dots & -\lambda & 0 \\ c_0 & 0 & \dots & 0 & -\lambda \end{vmatrix}$$

with

$$c_r = \sum_{i=0}^{\left[\frac{r}{2}\right]} a_{r-2i}\left(\frac{N-r+2i}{i}\right) \qquad (r = 0, 1, 2, \dots, n).$$

In the problem under consideration, which goes back to Tchebysheff [100], we talk about the best approximation of a given function

$$f(x) = a_0 x^N + \frac{a_1}{2} x^{N-1} + \dots + \frac{a_n}{2^n} x^{N-n}$$

in the interval $[-1, 1]$ by means of a rational fraction

$$R(x) = \frac{q_0 x^{N-1} + q_1 x^{N-2} + \dots + q_{N-1}}{p_0 x^n + \dots + p_n}.$$

By virtue of Tchebysheff's general theorem (cf. Section 34), the extremal fraction, which (after reduction) assumes its most simple form

$$R_0(x) = \frac{\beta_0 x^{N-\tau-1} + \dots + \beta_{N-\tau-1}}{\alpha_0 x^{n-\sigma} + \dots + \alpha_{n-\sigma}}$$

is characterized by the fact that the difference

$$y = f(x) - R_0(x)$$

attains its maximum value L with alternate change in sign at least $N + n - d + 1$ times in the interval $[-1, 1]$, where $d = \min\{\sigma, \tau\}$.

Let us consider the function

$$y^2 - L^2, \tag{8}$$

which has exactly $2(N + n - \sigma)$ zeros.

If the above-mentioned condition characterizing the function y is to be satisfied, all the zeros of the function (8) must be situated in the interval $[-1, 1]$. In this connection, $-1, 1$ must be simple zeros of this function; but the remaining $2(N + n - \sigma - 1)$ zeros must be equal in pairs, so that in the aggregate these constitute a system of $N + n - \sigma - 1$ double zeros situated in the interior of the interval $(-1, 1)$ (from which it is evident that $d = \sigma$). Now let us consider the function

$$\eta = y + \sqrt{y^2 - L^2}, \tag{9}$$

which is numerically equal to L in the interval $[-1, 1]$, since the inequality $-L \leq y \leq L$ holds in the interval $[-1, 1]$. Every branch of the function is single-valued in the x-plane, slit along the segment $(-1, 1)$. This follows from the fact that the multiplicity of every pole of the function (8) is even and that all the zeros, whose number is likewise even, are situated in the interval $[-1, 1]$. Let us take the branch which becomes infinite when $x = \infty$.

In order to map the slit x-plane upon the unit circle of the z-plane, let us set

$$x = \frac{1}{2}\left(z + \frac{1}{z}\right).$$

The quantity η, as a function of z, has the following properties: $\eta(z)$ is single-valued in the circle $|z| \leq 1$; $|\eta(z)|$ is equal to L along the circle $|z| = 1$; at the point $z = 0$, the function $\eta(z)$ has a pole whose principal part is

$$\frac{1}{2^{N-1}}\left\{\frac{c_0}{z^N} + \frac{c_1}{z^{N-1}} + \ldots + \frac{c_n}{z^{N-n}} + \ldots\right\},$$

where the numbers c_i are to be determined from the numbers a_i in the relations as given in the statement of our problem.

Let j ($j \leq n - \sigma$) denote the true number (= the number with due regard to multiplicity) of poles of the function $\eta(z)$ in the region $|z| < 1$, excluding the pole $z = 0$. It is clear that $n - \sigma - j$ is the exact number of zeros of the function $\eta(z)$ in the region $|z| < 1$. As the point z describes the circle $|z| = 1$, the argument of the function $\eta(z)$ changes by

$$2\pi(n - \sigma - j) - 2\pi j - 2\pi N.$$

Therefore the number of deviation points of the function y in the interval $[-1, 1]$ is not greater than

$$N + 2j - n + \sigma + 1.$$

Whence it follows that $j = n - \sigma$ and, moreover, that the function $\eta(z)$ has no zero in the region $|z| < 1$. If we apply Schwarz's principle of reflection, we can conclude from the foregoing that

$$\eta(z) = L\frac{1}{z^N}\frac{\bar{A}_{n-\sigma}z^{n-\sigma} + \ldots + \bar{A}_0}{A_0 z^{n-\sigma} + \ldots + A_{n-\sigma}} = \frac{1}{2^{N-1}}\left\{\frac{c_0}{z^N} + \frac{c_1}{z^{N-1}} + \ldots + \frac{c_n}{z^{N-n}} + \ldots\right\}.$$

Since all the poles of the function $\eta(z)$ lie in the region $|z| < 1$, we have by Section **19**:

$$L \cdot 2^{N-1} = \mu_n[F]$$

with

$$F(z) = \frac{c_0}{z^{n+1}} + \frac{c_1}{z^n} + \ldots + \frac{c_n}{z}.$$

Consequently, if we take into account Section **20**, the smallest positive root of the polynomial $D_n(\lambda)$ is $L \cdot 2^{N-1}$. It remains only to be observed that the numbers c_i are real and that therefore

$$D_n(\lambda) = (-1)^{n+1}\Delta_n(\lambda)\,\Delta_n(-\lambda)$$

where $\Delta_n(\lambda)$ is the determinant appearing in the statement of the theorem.

E. Solotareff's Problems and Related Problems

25. Solotareff's first problem [28]: Among all polynomials of the form

$$P(x) = x^n - n\sigma x^{n-1} + p_2 x^{n-2} + \ldots + p_n,$$

where σ is a given real number, it is required to find the particular polynomial that deviates the least from zero in the interval $[-1, 1]$.

This problem is a generalization of the problem of Tchebysheff considered in Section 36.

We solve it by a natural generalization [101] of the method used in Sections 37, 38. Moreover, we can limit ourselves to the case $\sigma \geq 0$ because, from the formula

$$F_2(x) = (-1)^n F_1(-x)$$

we obtain the solution $F_2(x)$ from the solution $F_1(x)$ for the case $\sigma = a$. Let y denote the desired solution and let L denote its deviation. By Tchebysheff's theorem (see Section 35), the polynomial y assumes the value $\pm L$ with alternate change in sign at no less that n points of the interval $[-1, 1]$.

The simplest polynomial of this kind is

$$y_1 = \frac{1}{2^{n-1}}\left(\frac{1+\beta}{2}\right)^n T_n\left(\frac{2x-\beta+1}{1+\beta}\right) = x^n - n\frac{\beta-1}{2}x^{n-1} + \dots$$

with

$$1 \leqq \beta \leqq 1 + 2\,\mathrm{tg}^2\frac{\pi}{2n}$$

and

$$T_n(x) = \cos n \arccos x.$$

The magnitude of the deviation is

$$L_1 = \frac{1}{2^{n-1}}\left(\frac{1+\beta}{2}\right)^n.$$

We have thus obtained a solution for

$$0 \leqq \sigma \leqq \mathrm{tg}^2\frac{\pi}{2n}.$$

The graphical representation of the polynomial y_1 is given schematically in Figures 6, 7, 8 for an odd n, at the two extreme values and at an intermediate value σ.

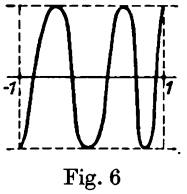


Fig. 6

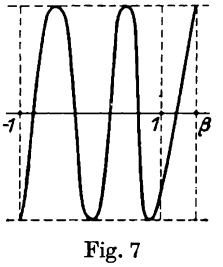


Fig. 7

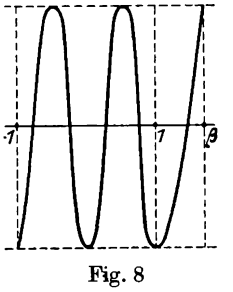


Fig. 8

When $\sigma > 0$ the number of deviation points is equal to n. If $\sigma \leq \mathrm{tg}^2\frac{\pi}{2n}$, the first one of the points

$$x_1 < x_2 < \dots < x_n$$

is a simple root, the rest are double roots of the equation

$$y^2 - L^2 = 0. \tag{1}$$

For $\sigma = \operatorname{tg}^2 \frac{\pi}{2n}$, we have

$$x_n = 1, \quad y(1) = -L, \quad y(\beta) = L.$$

Let us permit σ to increase, beginning with the value $\operatorname{tg}^2 \frac{\pi}{2n}$. Since the solution y varies continuously at the same time, the equations

$$x_1 = -1, \quad x_n = 1$$

will continue to hold. But x_n we know is a simple root of equation (1), hence we obtain an additional root α $(1 < \alpha < \beta)$ of this equation. By reason of the foregoing, the graphical representation of y will be as shown schematically in Figure 9.

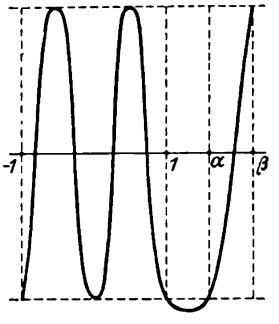


Fig. 9

Thus the solution y for $\sigma > \operatorname{tg}^2 \frac{\pi}{2n}$ has the following properties: Equation (1) has $n - 2$ double roots in the open interval $(-1, 1)$ and the simple roots $-1, 1, \alpha, \beta$, with $1 < \alpha < \beta$. Therefore, in the intervals $[-1, 1]$ and $[\alpha, \beta]$ we have the inequality

$$y^2 \leqq L^2,$$

and in the rest of the points on the real axis we have

$$y^2 > L^2.$$

To find a y with these properties, let G denote the complex x-plane, slit along

$$[-1, 1], \ [\alpha, \beta]$$

and let us consider the function

$$\eta = \frac{y + \sqrt{y^2 - L^2}}{L},$$

on said plane, where the root sign is so chosen that η becomes infinite with y. The function is:

1. numerically equal to 1 on the boundary of the region G;
2. different from zero in the interior of G and has a pole of the n-th order at infinity;
3. single-valued in the interior of G.

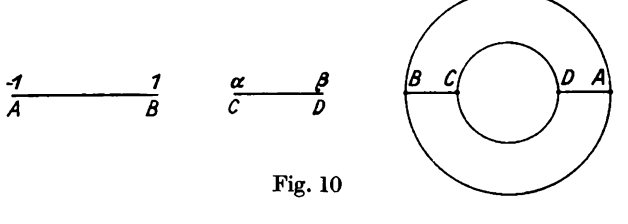


Fig. 10

Whence we infer that η can be expressed in terms of Green's function of a complex variable for the region G. For that purpose, let us map the region G upon a certain ring (denoted by Γ)

$$r < |v| < 1$$

of the v-plane, in a conformal manner, and so that the point at infinity of the region G passes into a certain (temporarily unknown) point $v = s$ of the positive real semi-axis of the v-plane. This mapping is given by the formulas

$$x = \frac{\mathrm{sn}^2 u + \mathrm{sn}^2 M}{\mathrm{sn}^2 u - \mathrm{sn}^2 M}, \qquad u = \frac{K' \ln v}{\pi}, \qquad M = \frac{K' \ln s}{\pi} \tag{2}$$

where the modulus $\varkappa$ of the elliptic function is

$$\varkappa = \sqrt{\frac{\alpha - 1}{\alpha + 1} \cdot \frac{\beta + 1}{\beta - 1}} \tag{3}$$

and the radius r is

$$r = e^{-\frac{\pi K}{K'}}.$$

The (temporarily unknown) parameters α, β are connected with the parameters $\varkappa$, M of formula (3) and we have

$$\mathrm{sn}^2 M = \frac{\beta - 1}{\beta + 1}, \tag{4}$$

which is equivalent to

$$\alpha = \frac{1 + \varkappa^2 \mathrm{sn}^2 M}{\mathrm{dn}^2 M}, \qquad \beta = \frac{1 + \mathrm{sn}^2 M}{\mathrm{cn}^2 M}. \tag{5}$$

The corresponding boundaries for this mapping are represented in Figure 10.

Green's function of a complex variable for the ring Γ with respect to the point s has the form [102]

$$\frac{H(u - M)}{H(u + M)}$$

with

$$H(u) = \vartheta_1\left(\frac{u}{2K}\right) = 2\sqrt[4]{q}\sin\frac{\pi u}{2K} - 2\sqrt[4]{q^9}\sin\frac{3\pi u}{2K} + \ldots$$

and

$$q = e^{-\frac{\pi K'}{K}}$$

Hence the function η, by reason of the properties 1, 2, has the form

$$\eta = \mu\left[\frac{H(u + M)}{H(u - M)}\right]^n;$$

here the constant μ is numerically equal to unity.

Since η is single-valued in Γ, we have

$$\eta(u + 2iK') = \eta(u).$$

But, as is well known,

$$H(u + 2iK') = e^{-\pi i e^{\frac{\pi K'}{K}}} e^{-\frac{\pi i u}{K}} H(u).$$

Consequently

$$M = \frac{mK}{n}$$

where m is an integer.

Inasmuch as y is unaltered by the substitution of $-u$ for u, it follows that $\mu = \pm 1$.

In order to determine m, we must remember that y assumes the value $\pm L$ with alternate change in sign exactly n times in the interval $[-1, 1]$. But if x varies from -1 to 1, the quantity u varies from 0 to iK'. The argument of the function

$$\left[\frac{H(M+u)}{H(M-u)}\right]^n$$

equals zero when $n = 0$ and must be equal to $\pm(n-1)\pi$ when $u = iK'$.

But since

$$\left[\frac{H(M+iK')}{H(M-iK')}\right]^n = e^{-\pi i n\left(1+\frac{M}{K}\right)}$$

we obtain for M the value

$$M = -\frac{K}{n},$$

and consequently, when

$$\sigma > tg^2\frac{\pi}{2n},$$

the solution is given by the formulas

$$y_2 = \frac{L_2}{2}\left\{\left[\frac{H\left(\frac{K}{n}-u\right)}{H\left(\frac{K}{n}+u\right)}\right]^n + \left[\frac{H\left(\frac{iK}{n}+u\right)}{H\left(\frac{K}{n}-u\right)}\right]^n\right\}$$

$$x = \frac{\mathrm{sn}^2 u + \mathrm{sn}^2\frac{K}{n}}{\mathrm{sn}^2 u - \mathrm{sn}^2\frac{K}{n}}.$$

But since the leading coefficient of the polynomial y_2 is 1, we have [6]

$$L_2 = \frac{1}{2^{n-1}}\left[\frac{\sqrt{\varkappa}\,\Theta_1^2(0)}{H_1\left(\frac{K}{n}\right)\Theta_1\left(\frac{K}{n}\right)}\right]^{2n}.$$

[6] The reader will recall the theta series

$$H_1(u) = \vartheta_2\left(\frac{u}{2K}\right) = 2\sqrt[4]{q}\cos\frac{\pi u}{2K} + 2\sqrt[4]{q^9}\cos\frac{3\pi u}{2K} + \ldots,$$

$$\Theta(u) = \vartheta_0\left(\frac{u}{2K}\right) = 1 - 2q\cos\frac{\pi u}{K} + 2q^4\cos\frac{2\pi u}{K} - \ldots,$$

$$\Theta_1(u) = \vartheta_3\left(\frac{u}{2K}\right) = 1 + 2q\cos\frac{\pi u}{K} + 2q^4\cos\frac{2\pi u}{K} + \ldots.$$

Moreover, for σ we find the expression

$$\sigma = \frac{2 \operatorname{sn} \frac{K}{n}}{\operatorname{cn} \frac{K}{n} \operatorname{dn} \frac{K}{n}} \left\{ \frac{1}{\operatorname{sn} \frac{2K}{n}} - \frac{\Theta'\left(\frac{K}{n}\right)}{\Theta\left(\frac{K}{n}\right)} \right\} - 1 .$$

It can be verified immediately that this quantity increases from $\operatorname{tg}^2 \frac{\pi}{2n}$ to ∞ as x describes the interval $[0, 1]$. Hence a $\varkappa$ $(0 < \varkappa < 1)$ can be found for $\sigma > \operatorname{tg}^2 \frac{\pi}{2n}$, and thus the required polynomial is completely determined.

26. Given a positive proper fraction x. Among all the fractions

$$y = \frac{\varphi(x)}{\psi(x)},$$

where $\varphi(x)$, $\psi(x)$ are polynomials of the n-th degree, find the particular fraction which deviates least from the function

$$\operatorname{sign} x$$

in the intervals

$$[-1, -\varkappa], \ [\varkappa, 1] .$$

This problem is essentially the same as Solotareff's fourth problem [28].

It is easy to prove [103]: If

$$y = \frac{\varphi(x)}{\psi(x)}$$

is a solution of the problem and $\mathfrak{M}$ $(0 < \mathfrak{M} < 1)$ is the deviation, then

a) only one of the equations

$$\varphi(x) = 0, \ \psi(x) = 0$$

can have roots in the interval $(-\varkappa, \varkappa)$;

b) the function

$$(1 - \mathfrak{M}^2) \frac{\psi(x)}{\varphi(x)}$$

also is a solution of the problem.

By reason of a), b), we can restrict ourselves to the problem of finding the solution that does not become infinite in the interval $(-\varkappa, \varkappa)$. This solution is completely characterized by the fact that the difference

$$y - \operatorname{sign} x$$

assumes the value $\mathfrak{M}$, with alternate change in sign, at $2n + 2$ consecutive points of the set E, consisting of the intervals

$$[-1, -\varkappa], \ [\varkappa, 1] .$$

In parametric representation, the solution is given by the formulas

$$(1)\qquad \begin{cases} x = \varkappa \operatorname{sn}(u, \varkappa), \\ y = (1 - \mathfrak{M}) \operatorname{sn}\left(\dfrac{u}{M}, \lambda\right), \\ \lambda = \dfrac{1 - \mathfrak{M}}{1 + \mathfrak{M}}, \end{cases}$$

$$(2)\qquad L' = \frac{K'}{nM}, \quad L = \frac{K}{M}$$

Here K, L are complete elliptic integrals of the first kind for the moduli $\varkappa$, λ, and K', L' are such for the complementary moduli.

Equations (1), (2) show that the solution is related to those transformations of elliptic functions in which the second period is divisible by an integer (transformatio realis secunda). We will not stop to derive the formula; the verification of the solution is easy.

In fact, from formulas (1), (2) we have

$$y = \frac{2\lambda}{1+\lambda}\frac{x}{M\varkappa} \prod_{p=1}^{\left[\frac{n}{2}\right]} \frac{1 + \dfrac{x^2}{\varkappa^2 \operatorname{tn}^2\left(\dfrac{2p}{n}K', \varkappa\right)}}{1 + \dfrac{x^2}{\varkappa^2 \operatorname{tn}^2\left(\dfrac{2p-1}{n}K', \varkappa'\right)}},$$

and accordingly

$$y = \frac{\varphi(x)}{\psi(x)},$$

wherein for an odd value of n the degree of the numerator is n and the degree of the denominator is $n - 1$; but for an even value of n the degree of the denominator is n and the degree of the numerator is $n - 1$, because of the relation

$$\frac{1}{\operatorname{tn}(K', \varkappa')} = \frac{i}{\operatorname{sn}(iK', \varkappa)} = 0.$$

Let us set

$$u = K + iv,$$

where v describes the interval $[0, K]$, and let x describe the interval $[\varkappa, 1]$. Here,

$$y = (1 - \mathfrak{M}) \operatorname{sn}\left(\frac{u}{M}, \lambda\right) = (1 - \mathfrak{M}) \operatorname{sn}(L + iw, \lambda),$$

where w describes the interval $[0, nL']$. Hence y varies in the interval $[1 - \mathfrak{M}, 1 + \mathfrak{M}]$ and assumes alternately the values $1 - \mathfrak{M}$, $1 + \mathfrak{M}$ at $n + 1$ points

$$w = 0, L', 2L', \ldots, nL',$$

corresponding to

$$v = 0, \frac{K'}{n}, \frac{2K'}{n}, \ldots, \frac{nK'}{n}$$

The problem under consideration is equivalent to the problem of finding a rational fraction which has numerator and denominator of given degree and which so approximates the function $\frac{1}{\sqrt{x}}$ in the interval $[\varkappa^2, 1]$ that the numerical maximum of the relative error is a minimum. Tchebysheff also studied this problem [28]. An analogous problem was considered in Section 39.

27. Among all polynomials of the n-th degree with unity as the leading coefficient, it is required to find the particular polynomial that deviates least from zero in the intervals

$$[-1, -a], [a, 1].$$

If the degree of the polynomial is even ($n = 2\nu$), the problem is solvable by elementary means. The desired polynomial then has the form

$$y = L_{2\nu} \cos \nu \arccos \frac{2x^2 - 1 - a^2}{1 - a^2}$$

and the minimal deviation is equal to

$$L_{2\nu} = \frac{1}{2^{2\nu-1}} (1 - a^2)^\nu.$$

When the degree is odd ($n = 2\nu - 1$), the solution can be represented by elliptic functions [101]

$$x = \frac{a \operatorname{cn} u}{\sqrt{a^2 - \operatorname{sn}^2 u}} \qquad (x = 1 \text{ für } u = 0),$$

$$y = \frac{L_{2\nu-1}}{2} \left\{ \left[\frac{H\left(\frac{K}{2\nu-1} + u\right)}{H\left(\frac{K}{2\nu-1} - u\right)} \right]^{\nu-\frac{1}{2}} + \left[\frac{H\left(\frac{K}{2\nu-1} - u\right)}{H\left(\frac{K}{2\nu-1} + u\right)} \right]^{\nu-\frac{1}{2}} \right\},$$

where

$$\operatorname{sn} \frac{K}{2\nu - 1} = a$$

and the minimal deviation is equal to

$$L_{2\nu-1} = \frac{1}{2^{2\nu-2}} \left[\frac{\Theta(0)\, \Theta_1(0)}{\Theta\left(\frac{K}{2\nu-1}\right) \Theta_1\left(\frac{K}{2\nu-1}\right)} \right]^{2\nu-1}$$

It can be proved that for $\nu \to \infty$ we have

$$L_{2\nu-1} \sim \frac{1}{2^{2\nu-2}} (1 - a^2)^{\nu-\frac{1}{2}} \sqrt{\frac{1+a}{1-a}}.$$

28. Let $\mathfrak{M}$ be a bounded closed point set in the plane of the complex variable x and let

$$L_n = \min_{A_k} \max_{x \in M} |x^n + A_1 x^{n-1} + \ldots + A_n|.$$

Then

$$\lim_{n\to\infty} \sqrt[n]{L_n} = \tau = \tau(\mathfrak{M})$$

exists. The quantity $\tau(\mathfrak{M})$ is called the transfinite diameter of the set $\mathfrak{M}$ and is equal to

$$\lim_{n\to\infty} \sqrt[\binom{n}{2}]{V_n},$$

where V_n is the largest value the modulus of the Vandermonde determinant

$$V(x_1, x_2, \ldots, x_n)$$

assumes as the points $x_1, x_2, \ldots, x_n$ run through $\mathfrak{M}$.

This theorem is due to Fekete [104]. According to the results in the previous section, the transfinite diameter of the point set consisting of the intervals

$$[-1, -a], [a, 1]$$

is equal to

$$\frac{1}{2}\sqrt{1-a^2}.$$

29. Let E be a point set consisting of the two intervals

$$[-1, a], [\beta, 1] \quad (-1 < a < \beta < 1)$$

and let

$$x = \sqrt{\frac{2(\beta - a)}{(1-a)(1+\beta)}}$$
$$1 = 2\,\mathrm{sn}^2 \varrho = a \quad (-K < \varrho < 0).$$

Then the transfinite diameter of the set E is

$$\tau = \frac{1}{2}\left[\frac{\Theta(0)\,\Theta_1(0)}{\Theta(\varrho)\,\Theta_1(\varrho)}\right].$$

Furthermore, suppose that $s(x)$ is a given positive function continuous on the set E and that

$$\mathscr{L}_n[s(x)] = \min_{A_k} \max_{x\in E} |s(x)\,(x^n + A_1 x^{n-1} + \ldots + A_n)|.$$

Let us put

$$\gamma = a + 2\,\frac{\mathrm{sn}\,\varrho\,\mathrm{cn}\,\varrho\,\Theta'(\varrho)}{\mathrm{dn}\,\varrho\;\;\Theta(\varrho)}.$$

Then the following statements are valid [105]:

1. If the number

$$-\frac{\varrho}{K} = R$$

is irrational, then, corresponding to an arbitrary ω satisfying the inequality

$$-K < \omega \leqq 0$$

there exists an infinite sequence of subscripts n_i with

$$\mathscr{L}_{n_i}[s(x)] \sim 2\tau^{n_i} \frac{\Theta(\varrho+\omega)}{\Theta(\varrho-\omega)} e^{\frac{1}{\pi}\int_E \ln s(x) \frac{|x-\gamma|\,dx}{\sqrt{(1-x^2)(\alpha-x)(\beta-x)}}}.$$

2. If the number R is rational and in its lowest terms has the form

$$R = \frac{P}{Q}$$

then there exist Q quantities

$$\sigma^{(0)},\ \sigma^{(1)},\ \ldots,\ \sigma^{(Q-1)},$$

which satisfy the inequality

$$-K \leqq \sigma \leqq 0$$

and correspond to the residue classes mod Q so that, for every sequence of mutually congruent subscripts n_i (mod Q), the asymptotic relation

$$\mathscr{L}_n[s(x)] \sim 2\tau^n \frac{\Theta(\varrho+\sigma^{(\mu)})}{\Theta(\varrho-\sigma^{(\mu)})} e^{\frac{1}{\pi}\int_E \ln s(x) \frac{|x-\gamma|\,dx}{\sqrt{(1-x^2)(\alpha-x)(\beta-x)}}}$$

is valid.

The above theorem is a generalization of Bernstein's theorem, which was based on one interval (cf. Section **13**).

30. (Continuation.) If $(2m+2)\,\varrho \equiv 0 \pmod{K}$

then [106]:

$$\min_{p_i} \int_E |x^{2m+1} + p_1 x^{2m} + \ldots + p_{2m+1}|\,dx =$$

$$= 2 \min_{p_i} \max_{x \in E} \left| (x^{2m+1} + p_1 x^{2m} + \ldots + p_{2m+1}) \frac{\sqrt{(1-x^2)(a-x)(\beta-x)}}{x-\gamma} \right| =$$

$$= 2 \min_{p_i} \max_{x \in E} |x^{2m+2} + p_1 x^{2m+1} + \ldots + p_{2m+2}| = 4\tau^{2m+2}.$$

F. The Best Harmonic Approximation of the Simplest Analytic Functions

31. Let $J_n[f(x)]$ denote the error in the best approximation of the function $f(x)$ in the space $L(-1, 1)$ by means of polynomials of degree n.

To prove:

a) $$\mathbf{J}_n\left[\frac{1}{(a-x)^s}\right] = \frac{2}{\Gamma(s)\,\Gamma(1-\{s\})}\int_a^\infty (t-a)^{-\{s\}}\frac{d^{[s]}}{dt^{[s]}}\ln\frac{(t+\sqrt{t^2-1})^{n+2}+1}{(t+\sqrt{t^2-1})^{n+2}-1}\,dt,$$

$$\text{for } a>1,\ s>0;$$

b) $$\mathbf{J}_n[(a-x)^s] = \frac{2\,|\sin \pi s|}{\pi}\int_a^\infty (t-a)\ln\frac{(t+\sqrt{t^2-1})^{n+2}+1}{(t+\sqrt{t^2-1})^{n+2}-1}\,dt,$$

$$\text{for } a\geqq 1,\ n\geqq s>-1;$$

c) $$\mathbf{J}_n[\ln(a-x)] = 2\int_a^\infty \ln\frac{(t+\sqrt{t^2-1})^{n+2}+1}{(t+\sqrt{t^2-1})^{n+2}-1}\,dt. \qquad \text{for } a\geqq 1.$$

The reasoning in Section 51 can be used to prove this proposition because the derivatives of arbitrary order of our function $f(x)$ do not change sign in the interval $(-1, 1)$.

Besides, we must bear in mind the relation

$$\int_{-1}^{1}\frac{\operatorname{sign}\sin k\varphi}{u-x}\,dx = 2\ln\frac{(u+\sqrt{u^2-1})^k+1}{(u+\sqrt{u^2-1})^k-1}$$

where $u>1$ and $x=\cos\varphi$.

Let us prove c), for example. Let $b>a>1$. The derivative of any order of the function

$$\ln\frac{b-x}{a-x} = \int_a^b \frac{du}{u-x}$$

retains its sign in the interval $(-1, 1)$. Therefore

$$\mathbf{J}_n\left[\ln\frac{b-x}{a-x}\right] = \int_{-1}^{1}\ln\frac{b-x}{a-x}\operatorname{sign}\sin(n+2)\varphi\,dx = 2\int_a^b \ln\frac{(u+\sqrt{u^2-1})^{n+2}+1}{(u+\sqrt{u^2-1})^{n+2}-1}\,du.$$

But since

$$\left|\mathbf{J}_n[\ln(a-x)] - \mathbf{J}_n\left[\ln\frac{a-x}{b-x}\right]\right| \leqq \mathbf{J}_n[\ln(b-x)] = \mathbf{J}_n\left[\ln\frac{b-x}{b}\right] \leqq$$

$$\leqq \int_{-1}^{1}\left|\ln\frac{b-x}{b}\right|dx < \ln\frac{b+1}{b-1}$$

it follows that

$$\mathbf{J}_n[\ln(a-x)] = 2\int_a^\infty \ln\frac{(u+\sqrt{u^2-1})^{n+2}+1}{(u+\sqrt{u^2-1})^{n+2}-1}\,du.$$

It is easy to see that this result remains valid for $a = 1$.

32. (Continuation.) Let us keep the notation of the previous section. It is required to prove that we have

α) for $n \geq s > -1$:

$$\mathbf{J}_n[(1-x)^s] \geqq \frac{2|\sin \pi s|}{\pi} \frac{\Gamma(2s+2)}{2^{s-1}(n+2)^{2s+2}} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^{2s+3}},$$

β) for $n \to \infty$:

$$\mathbf{J}_n[(1-x)^s] \sim \frac{2|\sin \pi s|}{\pi} \frac{\Gamma(2s+2)}{2^{s-1}n^{2s+2}} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^{2s+3}} \qquad (s > -1),$$

$$\mathbf{J}_n[\ln(1-x)] \sim \frac{4}{n^2} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^3},$$

$$\mathbf{J}_n\left[\frac{1}{(a-x)^s}\right] \sim \frac{4n^{s-1}}{|\Gamma(s)|} \frac{(a-\sqrt{a^2-1})^{n+2}}{(a^2-1)^{\frac{s-1}{2}}} \qquad (a > 1, s \geqq 0),$$

$$\mathbf{J}_n[\ln(a-x)] \sim \frac{4\sqrt{a^2-1}}{n} (a-\sqrt{a^2-1})^{n+2} \qquad (a > 1).$$

33. Let $\mathbf{E}_n[f(x)]$ denote the error in the best approximation of the function $f(x)$ in the space $C(-1, 1)$ with the aid of polynomials of degree n. For $n \to \infty$ Bernstein [67] proved the validity of the following asymptotic formulas:

a) $$\mathbf{E}_n\left[\frac{1}{(a-x)^s}\right] \sim \frac{n^{s-1}}{|\Gamma(s)|} \frac{(a-\sqrt{a^2-1})^n}{(a^2-1)^{\frac{s-1}{2}}} \qquad (a > 1, s \geqq 0),$$

b) $$\mathbf{E}_n[\ln(a-x)] \sim \frac{1}{n} \frac{(a-\sqrt{a^2-1})^n}{\sqrt{a^2-1}} \qquad (a > 1).$$

The formula a) can be proved for a positive integer s by the methods of Sections 37, 38, where it was proven for $s = 1, 2$.

We will not stop for the proof of this formula for odd values of s, nor for the proof of the formula b); instead, we refer the reader to a monograph by S. N. Bernstein [68].

34. Given a complex number A, a positive number c and a positive integer n. Prove that the equation

$$\mu_n(A) = \min_{\substack{a_k, b_k \\ f \in Q_T}} \sup_{-\infty < x < \infty} \left| \frac{A}{(x-ci)^n} + \frac{\bar{A}}{(x+ci)^n} - \sum_{k=1}^{n-1} \frac{a_k x + b_k}{(x^2+c^2)^k} - \varphi(x) \right| =$$

$$= \frac{|A|}{2^{n-1}c^n} e^{-cT}$$

is true for all positive values of T.

Let us consider the function

$$y = \frac{L}{2}\left\{e^{-iTx-i\delta}\left(\frac{x+ic}{x-ic}\right)^n + e^{iTx+i\delta}\left(\frac{x-ic}{x+ic}\right)^n\right\},$$

where δ is a real-valued parameter and L a positive parameter. η is a meromorphic function with two poles of order n. On the real axis, η is real-valued and we have:

$$\max_{-\infty<x<\infty} |y| = L.$$

If the parameters L and δ are so determined from the condition

$$\frac{L}{2} e^{cT-i\delta}(2ic)^n = A$$

that

$$L = \frac{|A|}{2^{n-1}c^n} e^{-cT}$$

then y assumes the form

$$y = \frac{A}{(x-ci)^n} + \frac{\overline{A}}{(x+ci)^n} - \sum_{k=1}^{n-1} \frac{a_k x + \beta_k}{(x^2+c^2)^k} - \psi(x)$$

with $\psi(x) \in G_T$.

Thus we have proved

$$\mu_n(A) \leqq L = \frac{|A|}{2^{n-1}c^n} e^{-cT}.$$

To prove that the equality sign is valid in this relation, let us choose any function different from y, to wit

$$z = \frac{A}{(x-ci)^n} + \frac{\overline{A}}{(x+ci)^n} - \sum_{k=1}^{n-1} \frac{a_k x + b_k}{(x^2+c^2)^k} - \varphi(x)$$

with $\varphi(x) \in G_T$.

We may evidently assume that z is real-valued along the real axis. Let us set

$$M = \sup_{-\infty<x<\infty} |z|$$

and prove that $M > L$.

To that end, we consider the set $\mathfrak{M}$ of all those points ξ of the upper x half-plane that are singular points for the function

$$\sqrt{z^2 - M^2}.$$

In every finite portion of the half-plane there exists only a finite number of such points ξ. Every $\xi \in \mathfrak{M}$ is the initial point of one or more arcs λ; $z(x)$ is real-valued along these arcs and satisfies the inequality

$$-M \leqq z(x) \leqq M.$$

The other end point of each of the arcs in question is either again a point of $\mathfrak{M}$ or a point lying on the boundary of the region $\Im(x) > 0$.

Let us cut the upper half-plane along the arcs λ. The half-plane will then be subdivided into $m \geq 1$ regions. One of them will contain the point ic; this region will be denoted by D. In the region D the function

$$\frac{z + \sqrt{z^2 - M^2}}{M} = g(x)$$

is evidently single-valued. The square root will be so defined here that $g(ic) = \infty$. Moreover, as x approaches any finite portion of the border of the region D, the numerical value of the function $g(x)$ converges uniformly to unity.

Let us now consider the function

$$h(x) = e^{iTx}\left(\frac{x - ic}{x + ic}\right)^n g(x),$$

which is regular in the region D. Its numerical value converges to the limit ≤ 1 uniformly, as x approaches any finite part of the border of the region D.

If, on the other hand, the point at infinity is a boundary point for the region D, then the inequality

$$|h(x)| < C,$$

where C is a constant, is satisfied in a certain neighborhood (belonging to D) of this point, in any case. In fact, in a neighborhood of the point at infinity we have

$$|g(x)| \leqq A + B|\varphi(x)|,$$

where A, B are certain constants easily evaluated. But since $\varphi(x) \in B_T$ it follows that

$$|g(x)| \leqq A + B_1 e^{Tx''} \qquad (x = x' + ix'')$$

and accordingly

$$|h(x)| \leqq e^{-Tx''}(A + B_1 e^{Tx''}) = Ae^{-Tx''} + B_1.$$

We can now apply the principle of the maximum. Because of this principle we have everywhere in D

$$|h(x)| \leqq 1,$$

and if the equality sign is valid at a point in the interior, then

$$h(x) \equiv e^{ia}.$$

If we take the point ic, we obtain

$$\frac{e^{-cT}}{M}\frac{2|A|}{(2c)^n} < t$$

or

$$\frac{L}{M} < 1.$$

This proves the proposition.

35. Given, a complex number A, a positive number c and a positive integer n. Using the result of the previous section, prove that [7]

a) $$\mathcal{E}_T\left[\frac{A}{x-ci}+\frac{\bar{A}}{x+ci}\right]=\frac{|A|}{c}e^{-cT} \qquad (T>0);$$

b′) for $T\to\infty$ and $n\geq 2$,

$$\mathcal{E}_T\left[\frac{A}{(x-ci)^n}+\frac{\bar{A}}{(x+ci)^n}\right]\sim\frac{|A|}{c}\frac{T^{n-1}}{(n-1)!}e^{-cT}.$$

If we let α, β be two real numbers, the above relations can be written:

a′) $$\mathcal{E}_T\left[\frac{ax+\beta}{x^2+c^2}\right]=\frac{\sqrt{a^2c^2+\beta^2}}{2c^2}e^{-cT} \qquad (T>0);$$

b) for $T\to\infty$ and $n\geq 2$,

$$\mathcal{E}_T\left[\frac{ax+\beta}{(x^2+c^2)^n}\right]\sim\frac{\sqrt{a^2c^2+\beta^2}}{2^nc^{n+1}}\frac{T^{n-1}}{(n-1)!}e^{-cT}.$$

Formula a) was discovered by Bernstein [69]. It is a special case of the results in Section **34**.

For the proof of proposition b), let us consider the expression

$$\Theta=\frac{1}{2}e^{-iTx}\sum_{k=1}^{n}L_ke^{-i\delta_k}\left(\frac{x+ic}{x-ic}\right)^k,$$

where L_k is a positive parameter and δ_k is a real-valued parameter. Let these parameters be determined so that

$$\Theta=\frac{A}{(x-ic)^n}+\Phi(x)$$

where $\Phi(x)$ is an integral function. Thus it is necessary to solve the following system of equations:

$$\frac{1}{2}e^{cT}L_ne^{-i\delta_n}(2ic)^n=A,$$

$$\sum_{k=r}^{n}L_ke^{-i\delta_k}\frac{1}{(k-r)!}\frac{d^{k-r}}{dx^{k-r}}\left[e^{-iTx}(x+ic)^k\right]\Big|_{x=ic}=0$$

$$(r=n-1,\ n-2,\ \ldots,\ 1).$$

Let us set

$$x_k=L_ke^{-i\delta_k}(2ic)^k(-iT)^k \qquad (k=1,2,\ldots,n)$$

[7] The symbol $\mathcal{E}_T[f(x)]$ is defined in Section 85.

and bring this system into the form

$$x_n = 2Ae^{-cT}(-iT)^n,$$

$$\sum_{k=r}^{n} \frac{x_k}{(k-r)!}(1+\varepsilon_k^{(r)}) = 0 \qquad (r = n-1, n-2, \ldots, 1),$$

where $\varepsilon_k^{(r)} = O\left(\frac{1}{T}\right)$ for $T \to \infty$. As can be readily seen, the solution of this system has the form

$$x_k = \frac{(-1)^{n-k}x_n}{(n-k)!}\left[1 + O\left(\frac{1}{T}\right)\right] \qquad (k = 1, 2, \ldots, n-1).$$

Hence we find

$$L_k = \frac{|A|}{2^{k-1}c^k}\frac{T^{n-k}}{(n-k)!}e^{-cT}\left[1 + O\left(\frac{1}{T}\right)\right] \qquad (k = 1, 2, \ldots n-1).$$

$$L_n = \frac{|A|}{2^{n-1}c^n}e^{-cT}.$$

By setting

$$\frac{L_k}{2}\left\{e^{-iTx-i\vartheta_k}\left(\frac{x+ic}{x-ic}\right)^k + e^{iTx+i\vartheta_k}\left(\frac{x-ic}{x+ic}\right)^k\right\} =$$

$$= \frac{A_k}{(x-ci)^k} + \frac{\overline{A}_k}{(x+ci)^k} - \sum_{r=1}^{k}\frac{\alpha_r^{(k)}x + \beta_r^{(k)}}{(x^2+c^2)^r} = \psi_k(x) = f_k(x) - \psi_k(x)$$

$$(k = 1, 2, \ldots, n),$$

where $A_n = A$, we can write

$$\frac{A}{(x-ci)^n} + \frac{\overline{A}}{(x+ci)^n} - f_1(x) = \sum_{k=2}^{n} f_k(x).$$

Whence it follows that

$$\left|\mathscr{E}_T\left[\frac{A}{(x-ci)^n} + \frac{A}{(x+ci)^n}\right] - \mathscr{E}_T[f_1(x)]\right| \leqq \sum_{k=2}^{n}\mathscr{E}_T[f_k(x)],$$

and it remains only to remember that by the previous section we have:

$$\mathscr{E}_T[f_k(x)] = \mu_k(A_k) = L_k \qquad (k = 1, 2, \ldots, n).$$

NOTES

[1] For further details on metric spaces (and also on linear normalized spaces) see E. Hausdorff, *Mengenlehre*, 2nd ed., Berlin 1927; S. Banach, *Théorie des opérations linéaires* (Warsaw 1932); L. A. Lusternik, "Fundamental concepts of functional analysis," (Russian, Acad. of Sciences).

[2] O. Hölder, "Über einen Mittelwertsatz" (Göttinger Nachrichten, 1889); H. Minkowski, *Diophantische Approximation*, Leipzig 1907; F. Riesz, *Les systèmes d'équations linéaires à une infinité d'inconnues*, Paris 1913.

[3] See, e.g., de la Vallée-Poussin, *Cours d'analyse*, Volume 2, Chapter 3, 2nd ed., Paris 1912.

[4] For further details on Hilbert space, see A. I. Plessner, *Spectral theory of linear operations*.

[5] This example is due to S. N. Bernstein; see S. N. Bernstein, *On polynomials orthogonal in a finite interval*. (In Russian, Kharkov 1937).

[6] I. P. Gram, "Über die Entwicklung reeller Functionen in Reihen mittels der Methode der kleinsten Quadrate" (Journal für Mathematik, 94, 1883).

[7] R. Courant-D. Hilbert, *Methoden der mathematischen Physik*, Berlin 1931.

[8] This determinant was already familiar to Cauchy; see G. Pólya-G. Szegö, *Aufgaben und Lehrsätze der Analysis*, Berlin 1925.

[9] This proof is due to F. Riesz, "Lineare Operatoren im komplexen Hilbertschen Raume" (Acta Sci. Math., Szeged, 5, 1930).

[10] E. Schmidt, "Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener," (Mathematische Annalen, 63, 1907).

[11] H. Bohr, "Fastperiodische Functionen." Ergebnisse der Mathematik, Berlin 1932.

[12] K. Weierstrass, "Über die analytische Darstellbarkeit sogenannter willkürlicher Funktionen einer reellen Veränderlichen," (Berliner Berichte 1885).

[13] S. N. Bernstein, "Démonstration du théorème de Weierstrass fondée sur le calcul de probabilité," (Proceedings of the mathematical society of Kharkov, XIII, 1912). There are a great many proofs of Weierstrass' theorem. Some of them can be found in the following treatises: De la Vallée-Poussin, "Cours d'analyse" (cf. [3]); W. L. Gontcharov. "Theory of Interpolation and of Approximation of Functions" (Russian, 1934).

[14] De la Vallée-Poussin, see [3].

[15] See, e.g., J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Berlin 1932.

[16] F. Riesz, "Sur les ensemble des fonctions" (Comptes rendus de l'Académie des Sciences, Paris, 143, 1906); E. Fischer, "Sur la convergence en moyenne," (l.c., 144, 1907).

[17] Ch. Müntz, "Über den Approximationssatz von Weierstrass" (Schwarz-Festschrift 1914).

[18] As to criteria for the closure of vector manifolds in arbitrary linear normalized spaces, see Section 30.

[19] The substance of this Section is taken from the book by S. Kaczmarz-H. Steinhaus, *Theorie der Orthogonalreihen*, Warsaw, 1935.

[20] See S. Banach [1] and L. A. Lusternik [1]. It is to be noted that in these works only real spaces are considered, that is to say the vectors are multiplied only by real numbers and the functionals assume only real values. The passage to the general case can, however, be accomplished without difficulty and was given by several authors simultaneously; see, for example, G. Suchomlinoff, *On the extension of linear functionals in spaces of a complex variable and in quaternion spaces* (in Russian, Mathematical Collection, new series, 1938).

Let G be a certain subspace of an arbitrary linear normalized space E, and let $\varphi(x)$ be a linear functional in G with the norm $||\varphi||_G = M$. For any $x \in G$ we have

$$\varphi(x) = \varphi_1(x) + i\,\varphi_2(x), \tag{1}$$

where $\varphi_1(x)$ and $\varphi_2(x)$ are real-valued linear functionals in G; in other words, for $x, y \in G$ and for real values of α, β we have

$$\varphi_k(\alpha x + \beta y) = \alpha\,\varphi_k(x) + \beta\,\varphi_k(y) \quad (k = 1, 2)$$

and

$$||\varphi_k||_G \leqq M. \tag{2}$$

Since the equation

$$\varphi(ix) = \varphi_1(ix) + i\varphi_2(ix)$$

subsists simultaneously with (1) and since, on the other hand, we have from (1)

$$\varphi(ix) = i\varphi(x) = -\varphi_2(x) + i\varphi_1(x),$$

it follows that the two equations

$$\varphi_1(ix) = -\varphi_2(x), \quad \varphi_2(ix) = \varphi_1(x)$$

hold. We see that the extension of the functional $\varphi(x)$ results in an extension of the real-valued functional $\varphi_1(x)$. Suppose this extension has been carried out so that we get

$$||\varphi_1||_E = ||\varphi_1||_G .$$

The extension of the functional $\varphi(x)$ to E will then be given by the formula

$$\varphi(x) = \varphi_1(x) - i\varphi_1(ix).$$

Let $x \in E$. If we set

$$y = \frac{\overline{\varphi(x)}}{|\varphi(x)|}\,x$$

we have

$$\varphi(y) = |\varphi(x)| = \varphi_1(y)$$

and therefore

$$|\varphi(x)| \leqq ||\varphi_1||_E \cdot ||y|| = ||\varphi_1||_G \cdot ||x||.$$

Consequently the inequality

$$||\varphi||_E \leqq ||\varphi_1||_G \leqq M$$

holds; and, since

$$||\varphi||_E \geqq ||\varphi||_G = M$$

it follows that

$$||\varphi||_E = ||\varphi||_G.$$

Accordingly, the functional $\varphi(x)$ has been extended without increasing the norm. (At the same time we have proved that the equality sign in (2) stands.)

[21] See F. Riesz [9]. For separable Hilbert spaces, the theorem was proved considerably earlier; see F. Riesz [2].

[22] See F. Riesz [2], S. Banach [1], L. A. Lusternik [1].

[23] See S. Banach [1].

[24] De la Vallée-Poussin, "Sur les polynômes d'approximation et la représentation approchée d'un angle," (Bulletin de l'académie belgique, 1911). In that paper, approximations through polynomials and trigonometric functions are considered.

As to the generalization of de la Vallée-Poussin's theorem, considered in Section 32, see N. I. Achieser, "On extremal properties of certain rational functions." (In Russian, Reports of the Soviet Academy of Sciences, 1930, pp. 495—499).

A generalization of de la Vallée-Poussin's theorem in another direction has been given by Bernstein. See, for example, S. N. Bernstein, *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle*, Paris.

[25] P. L. Tchebysheff, *Problems on minimum expressions connected with the approximate representation of functions.* (In Russian, Collected Papers, Vol. I, 1899).

The problems considered in Chapter 2 of our text are related to investigations made by Tchebysheff and described by him in his papers.

[26] P. L. Tchebysheff, *On polynomials which yield the best values for the simplest rational functions.* (In Russian, Collected Papers, Vol. I).

See also Achieser, "On certain functions with the least deviation from zero in given intervals", (Proceedings of the Society of Mathematical Physics of Kazan, 1928.)

[27] See S. N. Bernstein [24], as well as N. I. Achieser, "On the asymptotic expressions of the best approximation of certain rational functions through polynomials," (Proceedings of the Kharkov Mathematical Society, 1932).

[28] E. I. Solotareff, Application of elliptic functions to problems concerning functions that deviate least from zero. N. I. Achieser, "On an extremal property of rational functions," (Proceedings of the Mathematical Society of Kharkov, 1933). N. I. Achieser, "Remarks on extremal properties of certain fractions," (Proceedings of the Mathematical Society of Kharkov, 1935).

[29] W. Cauer, "Ein Interpolationsproblem für Funktionen mit positivem Realteil," (Mathematische Zeitschrift, 38, 1934).

[30] This example is due to Bernstein; see S. N. Bernstein [24].

[31] A. Haar, "Die Minkowskische Geometrie und die Annäherung an stetige Funktionen" (Mathematische Annalen, 78, 1918).

In the proof of the sufficiency of his condition, Haar employs the theory of convex bodies. The proof given in Section 45 is based on the general reasoning of Tchebysheff and A. Markoff.

[32] The connection between the problem in Section 46 and Haar's problem was called to the attention of the auth or by Krein. As to other problems connected with the systems of Section 46, see the paper by Krein, "The L-problem in abstract normalized spaces." N. I. Achieser and M. G. Krein, On some problems of the theory of moments. See also E. Remes, "On the best representation of functions in the Tchebysheff sense" (Ukranian; Kiev, 1935).

[33] See S. N. Bernstein [24].

[34] D. Jackson, "A general class of problems in approximation," (American Journal of Mathematics, XLVI, 1924, pp. 215—234).

Jackson considers the case where $f_1(x) = x^{i-1}$ $(i = 1, \ldots, n)$. The proof in the text is a slight modification of Jackson's. A different proof as well as other interesting theorems relating to this subject will be found in the work of Krein referred to in [32].

[35] P. L. Tchebysheff, "On interpolation in the case of a large number of experimental data." (Russian, Papers, Vol. 1, 1859).

[36] A. A. Markoff, "On the asymptotic values of integrals, in relation to interpolation." (Russian, Papers of the Academy of Sciences, Series VIII, Vol. VI, 1898).

[37] A. N. Korkin and E. I. Solotareff, "Sur un certain minimum" (Collected works of Solotareff, Vol. 1).

[38] A. Zygmund, *Trigonometrical series* (Warsaw, 1935).

[39] W. H. Young, "On the integration of Fourier Series," (Proceedings of the London Mathematical Society, 9, 1911).

[40] N. E. Nörlund, *Vorlesungen über Differenzenrechnung* (Berlin 1924); de la Vallée-Poussin [3].

[41] E. C. Titchmarsh, *Introduction to the theory of Fourier integrals* (Oxford 1937).

[42] M. Plancherel, "Contribution à l'étude de la représentation d'une fonction arbitraire par des intégrales définies," (Rendiconti del Circolo Matematico di Palermo, 1910).

[43] G. N. Watson, "General transforms" (Proceedings of the London Mathematical Society, 1933).

[44] E. C. Titchmarsh, "A proof of a theorem of Watson" (Journal of the London Mathematical Society, 8, 1933).

[45] L. Fejér, *Untersuchungen über Fouriersche Reihen* (Mathematische Annalen, 58, 1904).

[46] De la Vallée-Poussin, "Leçons sur l'approximation des fonctions d'une variable réelle" (Paris, 1919).

[47] S. Bochner, *Vorlesungen über Fouriersche Integrale* (Leipzig, 1932); E. W. Hobson, *The theory of functions of a real variable and the theory of Fourier's series*, 2nd ed. (Cambridge 1926); E. C. Titchmarsh [41].

[48] H. Lebesgue, *Leçons sur les séries trigonométriques* (Paris, 1906).

[49] G. H. Hardy, "Notes on some points in the integral calculus," (Messenger of Mathematics, 1922).

[50] A. G. Offord, "On the uniqueness of the representation of a function by a trigonometric integral" (Proceedings, London Mathematical Society, 52, 1937).

[51] A. G. Offord, "Note on the uniqueness of the representation of a function by a trigonometric integral" (Journal, London Mathematical Society, 11, 1936).

[52] The class of functions W^2 plays an important part in the generalized harmonic analysis of Wiener. See N. Wiener, *The Fourier integral and certain of its applications* (Cambridge, 1933); S. Bochner [47].

[53] E. Borel, *Leçons sur les séries divergentes* (Paris, 1901). Actually, the Borel transformation is nothing else but the Laplace transformation. For further details concerning the latter, see Widder, *The Laplace Transform* (New York, 1941).

[54] R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain* (New York, 1934).

[55] H. Plancherel and G. Pólya, "Fonctions entières et intégrales de Fourier," (Comm. Helv. 9, 1937).

[56] S. N. Bernstein, "Sur une propriété des fonctions entières" (Comptes rendus 176, 1923).

[57] S. N. Bernstein, "On the best approximation of continuous functions through polynomials of given degree," (Communications, Kharkov, Mathematical Society, 1912).

It is to be noted that long before the work of Bernstein, A. Markoff and W. Markoff had proved inequalities closely related to Bernstein's for trigonometric functions. See A. Markoff, "On a problem of Mendeleyev," (Communications, Academy of Sciences, 1890). W. Markoff, "On functions which deviate least from zero in a given interval." (Communications, Kharkov, 1892.)

W. Markoff's theorem reads: If $P_n(x)$ is a polynomial of the n-th degree, then we have

$$\max_{-1 \leqq x \leqq 1} |F_n^{(k)}(x)| \leqq 2^k \frac{n^2(n^2-1^2)\dots(n^2-[k-1]^2)}{(k+1)\dots 2k} \max_{-1 \leqq x \leqq 1} |P_n(x)|.$$

A. Markoff investigated the case $k = 1$.

Of the various proofs of Bernstein's inequality, the one by Riesz is of particular interest. See Riesz, "Formule d'interpolation pour la dérivée d'un polynôme trigonométrique," (Comptes rendus (158, 1914); "Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome," (Jahresberichte der Deutschen Mathematischen Vereinigung, 1914).

[58] G. Szegö, "Über einen Satz des Herrn Bernstein," (Schriften der Königlichen Gelehrten Gesellschaft, 1928); see also S. N. Bernstein, "Sur un théorème de M. Szegö" (Prace Matem. Fizyczne, 42, 1935).

[59] M. G. Krein, "On the representation of a function through Fourier-Stieltjes integrals," (Kuybishev, 1943).

[60] B. M. Levitan, "On a generalization of the inequalities of S. N. Bernstein and H. Bohr," (Russian, Papers of the Soviet Academy of Sciences, XV, 1937).

[61] See Fejér, "Über trigonometrische Polynome" (Journal für Mathematik 146, 1916), and also Pólya-Szegö [8].

[62] Bochner, "A theorem on Fourier-Stieltjes integrals," (Bulletin of the American Mathematical Society, XL, 1934).

[63] Lebesgue, "Sur la représentation approchée des fonctions," (Rendiconti del Circolo Matematico di Palermo, 26, 1908).

[64] De la Vallée-Poussin, "Note sur l'approximation par un polynôme d'une fonction dont la derivée est à variation bornée," (Bulletin de l'académie Belgique, 1908).

[65] See S. N. Bernstein [57]. This distinguished work was awarded a prize by the Belgian Academy of Sciences. In his review, de la Vallée-Poussin characterized Bernstein's work as "the most significant contribution ever made to the field of the decomposition of real functions in series of polynomials, — the most significant as to both the number and the importance of the results contained therein."

Among the papers on this subject, the following deserve attention: de la Vallée-Poussin [24]; S. N. Bernstein, "Sur l'approximation des fonctions continues par des polynômes," (Comptes rendus de l'Académie des Sciences, Paris, 152, 1911).

[66] D. Jackson, "Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen gegebenen Grades und trigonometrische Summen gegebener Ordnung," (Dissertation, Göttingen, 1911). See also the later monograph: D. Jackson, *The Theory of Approximation* (New York, 1930).

[67] S. N. Bernstein, "Sur la valeur asymptotique de la meilleure approximation des fonctions analytiques admettant des singularités données," (Bulletin de l'Académie Belgique, 1913).

[68] S. N. Bernstein, *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytique d'une variable réelle*, (Paris, 1926). See also S. N. Bernstein [24].

[69] S. N. Bernstein, "Constructive function theory as a development of Tchebysheff's ideas," (Russian, Communications of the Soviet Academy of Sciences, 9, 1945).

[70] G. H. Hardy, "Weierstrass' non-differentiable function" (Transactions of the American Mathematical Society, 17, 1916). A. Zygmund, "Smooth functions," (Duke Mathematical Journal, 12, 1945).

[71] J. Favard, "Sur les meilleure procédés d'approximation" (Bulletin de la Société Mathématique de France, LXI, 1937); N. Achieser and M. Krein, "On the best approximation of periodic functions" (Russian, Communications of the Soviet Academy of Sciences, XV, 1937). N. Achieser, "On the best approximation of a class of continuous periodic functions," (l.c. XVII, 1937). M. Krein, "On the theory of the best approximation of periodic functions," (l.c., XVIII, 1938).

The first two papers deal with the class of functions $f(x)$ with the period 2π where

$$|f^{(r)}(x)| \leqq 1 \ .$$

The other two papers deal with the class of periodic functions for which the inequality

$$\left| P\left(\frac{d}{dx}\right) f(x) \right| \leqq 1$$

is valid, where

$$P(D) = D^r + A_1 D^{r-1} + \ldots + A_r$$

is a given polynomial subject to one or the other of these restrictions.

[72] M. G. Krein, "On the best approximation of functions which are continuous and differentiable along the entire real axis," (Communications of the Soviet Academy of Sciences, XVIII, 1938).

[73] B. Sz.-Nagy, "Über gewisse Extremalfragen bei transformierten trigonometrischen Entwicklungen." I. Periodischer Fall (Leipziger Berichte, XC, 1938). II. Nichtperiodischer Fall (Leipziger Berichte, XCI, 1939).

[74] Special cases of the inequality for $f(x)$ are due to H. Bohr ($r = 1$, $n \geq 1$) and S. N. Bernstein ($r \geq 1$, $n = 1$). The general case was given by J. Favard [71] and for $\tilde{f}(x)$ by N. I. Achieser and M. G. Krein [71]. Interesting generalizations are due to B. M. Levitan [60].

See H. Bohr, "Un théorème général sur l'intégration d'un polynôme trigonométrique," (Comptes rendus, 200, 1935). S. B. Bernstein, "Sur quelques propriétés extrémales des intégrales successives," (Comptes rendus, 200, 1935).

[75] Jackson proved the theorems relative to the approximation of differentiable periodic functions through trigonometric sums by considering the integral

$$h_n \int_{-\pi}^{\pi} f(x+t) \left[\frac{\sin \frac{nt}{2}}{n \sin \frac{t}{2}} \right]^4 dt$$

with

$$h_n = \frac{1}{\int_{-\pi}^{\pi} \left[\frac{\sin \frac{nt}{2}}{n \sin \frac{t}{2}} \right]^4 dt} \ .$$

which is analogous to the Fejér integral.

De la Vallée-Poussin introduced a similar integral, namely

$$I_n(x) = \frac{3}{2\pi} \int_{-\infty}^{\infty} f\left(x + \frac{2t}{n}\right) \frac{\sin^4 t}{t^4} dt,$$

which we have already mentioned in Section 64. To this integral is attached the "convergence factor"

$$D(a) = \begin{cases} 1 - \frac{3}{2}a^2 + \frac{3}{4}a^3 & (0 \leqq a \leqq 1), \\ \frac{1}{4}(2-a)^3 & (1 \leqq a \leqq 2), \\ 0 & (a \geqq 2). \end{cases}$$

If $f(x)$ is a periodic function with the continuity modulus $\omega(\delta)$, then

$$|f(x) - I_n(x)| \equiv \left| f(x) - \frac{a_0}{2} - \sum_{k=1}^{2n-1} D\left(\frac{k}{n}\right)(a_k \cos kx + b_k \sin kx) \right| \leqq 2\omega\left(\frac{\pi}{2n}\right),$$

where

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

is the Fourier series of the function $f(x)$.

The proof is readily obtained from the fact that

$$I_n(x) - f(x) = \frac{3}{2\pi} \int_0^{\infty} \left\{ f\left(x + \frac{2t}{n}\right) + f\left(x - \frac{2t}{n}\right) - 2f(x) \right\} \frac{\sin^4 t}{t^4}\, dt .$$

This representation enables us to obtain for

$$|f(x) - I_n(x)|$$

an estimate in which, instead of the continuity modulus $\omega(\delta)$, there appears the quantity $\omega^*(\delta)$; the definition of $\omega^*(\delta)$ has been given in Section 79. In particular the inequality

$$|f(x) - I_n(x)| \leqq \text{const} \cdot \frac{1}{n}$$

is valid, if $f(x)$ satisfies the condition A^* besides having $f(x) \in \text{Lip } 1$. In connection therewith see A. Zygmund [70].

Let us set

$$I_{n,r}(x) = \frac{a_0}{2} + \sum_{k=1}^{2n-1} \left\{ 1 - \left[1 - D\left(\frac{k}{n}\right) \right]^r \right\} (a_k \cos kx + b_k \sin kx).$$

If $f(x)$ has a continuous derivative of order $r \geq 1$, then

$$|f(x) - I_{n,r}(x)| \leqq 2\left(\frac{\pi}{2n}\right)^r \omega\left(\frac{\pi}{2n}; f^{(r)}\right).$$

[76] N. I. Achieser and B. M. Levitan, "On an application of the inequality of H. Bohr and J. Favard." (Communications of the Soviet Academy of Sciences, XIV, 1937). See also de la Vallee-Poussin [24].

[77] A. N. Kolmogoroff proved that if

$$|f^{(r)}(x)| \leqq 1,$$

then the inequality

$$\sup_f |f(x) - s_{n-1}(x)| = \frac{4}{\pi^2} \frac{\ln n}{n^r} (1 + \varepsilon_n)$$

is valid, where $\varepsilon_n \to 0$ as $n \to \infty$. See A. N. Kolmogoroff, "Zur Größenordnung des Restgliedes Fourierscher Reihen differenzierbarer Funktionen," (Annals of Mathematics, 36, 1935).

Further results of this kind are due to S. M. Nikolski. See Nikolski, "The approximation of periodic functions through trigonometric polynomials," (Russian, Papers of the Steklov Mathematical Institute, XV, 1945).

[78] I. I. Priwaloff, "Sur les fonctions conjugées" (Bulletin de la Société Mathématique de France, 44, 1916); see also A. Zygmund [70].

[79] E. S. Quade, "Trigonometric Approximation in the Mean," (Duke Journal Math. Journal, 3, 1937).

[80] N. I. Achieser, "On the best approximation of analytic functions," (Communications of the Soviet Academy of Sciences, XVIII, 1938). This pertains to problems of approximation of harmonic functions. These problems have been treated in the works of Krein and Sz.-Nagy. See M. G. Krein [71] and B. Sz.-Nagy [73]. In particular these authors have established the following facts: Let $h(r, \varphi)$ be a harmonic function in the circle $r < 1$, such that

$$|h(r, \varphi)| \leqq 1 \qquad (0 \leqq \varphi \leqq 2\pi) ;$$

if we set

$$h(r, \varphi) = \sum_{k=n}^{\infty} r^k (a_k \cos k\varphi + b_k \sin k\varphi) ,$$

so that

$$\tilde{h}(r, \varphi) = \sum_{k=n}^{\infty} r^k (b_k \cos k\varphi - a_k \sin k\varphi) ,$$

then it follows that

$$|h(r, \varphi)| \leqq \frac{4}{\pi} \operatorname{arctg} r^n,$$

$$|\tilde{h}(r, \varphi)| \leqq \frac{2}{\pi} \ln \frac{1 + r^n}{1 - r^n}.$$

For $n = 1$ the first of these inequalities was found by H. A. Schwarz, the second by P. Koebe. See H. A. Schwarz, Gesammelte Abhandlungen, 2 (Berlin, 1890); P. Koebe, "Über das Schwarzsche Lemma und einige damit zusammenhängende Ungleichheitsbeziehungen," (Mathematische Zeitschrift, 6, 1920).

[81] See N. Wiener [52], also Widder [53]. A number of simplifications in the structure of Wiener's work have been introduced by W. A. Ditkin. In our presentation we have made use of these. See W. A. Ditkin, "The investigation of the structure of ideals in certain normalized rings," (Russian, 1939).

[82] See N. Wiener [52], Paley-Wiener [54], P. Lévy, "Sur la convergence absolue des séries de Fourier," (Comptes Mathematiques 1, 1934).

[83] See Widder [53].

[84] A. Tauber, "Ein Satz aus der Theorie der unendlichen Reihen," (Monatshefte für Mathematik und Physik, 8, 1897).
[85] S. Ikehara, "An extension of Landau's theorem in the analytic theory of numbers," (Journal of Mathematics and Physics, Massachusetts Institute of Technology, 10, 1931); see also Wiener [52] and Widder [53].
[86] T. Carleman, "Propriétés asymptotiques des fonctions fondamentales des membranes vibrantes," (Scandinavian mathematical congress, 1934).
[87] S. N. Bernstein, "Sur une classe des polynômes orthogonaux," (Communications of the Mathematical Society, Kharkov, Series IV, Vols. IV, V). See also Bernstein [5].
[88] A. A. Markoff, "The determination of certain functions from the condition that they deviate least from zero." (Communications of the Mathematical Society, Kharkov, 1884).
[89] See B. G. Szegö, *Orthogonal Polynomials* (New York, 1939).
[90] A. N. Kolmogoroff, "Stationary sequences in Hilbert space," (Russian, 1941).
[91] M. G. Krein, "On a generalization of the investigations of G. Szegö, W. M. Smirnoff and A. N. Kolmogoroff" (Communications of the Soviet Academy of Sciences, XLVI, 1945).
[92] See, for example, I. I. Priwaloff, "Properties of the limit value of single-valued analytic functions," (Russian, 1941).
[93] M. G. Krein, "On an interpolation problem of A. N. Kolmogoroff" (Communications of the Soviet Academy of Sciences, XLVI, 1945).
[94] See, for example, I. M. Vinogradoff, "Foundations of the theory of numbers," (Russian, 1944).
[95] O. Szász, "Über Approximation stetiger Funktionen durch Bernoullische Polynome," (Journal für Mathematik, Bd. 148).
[96] N. I. Achieser, "On a minimum problem of function theory and on the number of roots of an algebraic equation which lie at the interior of the unit circle" (Communications of the Academy of Sciences, 1931). See also the correction of this article by Krein and Neimark, "The method of symmetrical and Hermitian forms in the theory of the separation of roots of algebraic equations" (Kharkov, 1936).
[97] Carathéodory and Fejér, "Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landauschen Satz," (Rendiconti, Palermo, 32, 1911). L. Fejér, "Über gewisse Minimumprobleme der Funktionentheorie," (Mathematische Annalen, 97, 1911).
[98] I. Schur, "Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind," (Journal für Mathematik, 147, 148; 1918).
[99] A. Cohn, "Über die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise," (Mathematische Zeitschrift, 14, 1922).
[100] See Tchebysheff [25], also Achieser, "Über ein Tschebyscheffsches Extremumproblem," (Mathematische Annalen, Bd. 104, 1931).
[101] See Achieser [26].
[102] See, e.g., H. Weber, "Die partiellen Differentialgleichungen der mathematischen Physik." I (1900), p. 354.
[103] N. I. Achieser, "On a problem of E. I. Solotareff," (Communications of the Soviet Academy of Sciences, 1929).
[104] M. Fekete, "Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten," (Mathematische Zeitschrift, 17, 1934).
[105] N. I. Achieser, "On some functions which deviate least from zero in two given intervals," (Communications of the Soviet Academy of Sciences, 1932—1933).
[106] N. I. Achieser, "Generalization of a minimum problem of Korkine and Solotareff," (Communications of the Mathematical Society, Kharkov, Series 4, XIII, 1936).

INDEX

Approximation
  of analytic functions 214, 303
  of continuous functions 76, 138
  of differentiable functions 187
  harmonic 160, 176, 303
  by periodic functions 66
  polynomials of the best 78
  in the space L^p 12
  in Hilbert space 15, 19
Approximation theory
  fundamental theorem of 10
  principal problem of 1
Asymptotic formulas
  of Achieser 298
  of Bernstein 291

Banach 47
Bernoulli's polynomials 100, 269
Biorthogonal systems 83
Bochner 299, 300
Bohr's almost periodic functions 29, 301
Borel transform 133, 135, 300
Borel transformation 132
Buniakovski 6

Carathéodory-Fejér
  extremal function of 275
  problem of 270, 274
Carleman's Tauber theorem 241
Cauer 64
Cauchy's test 14
Class
  B_σ 137
  E_σ 132
  W_σ 134
Closure criteria
  for a set of vectors 49
Cluster point 14
Compactness 139
Conjugate
  function 126
  quantity 8
  series 126
Convergence
  of a sequence 14
  in normal form (in the norm) 40
  in the mean 107
Convoluted function 105, 122

De la Vallée Poussin
  problem of 160, 302
Deviation points 55
Dini-Lipschitz condition 162
Distance between two points 1
Ditkin 304

Eigenvalues 242, 276, 278
Elliptic integrals 62
Elliptic functions 62
Equicontinuous functions 139
Error, square of 16
Extremal function of Carathéodory 275
Extremal polynomial 250
Extremal problem 243

Faltung of functions 105, 122
Fatou's lemma 261
Favard 301, 303
Fejér's
  integral 112, 202
  kernel 111
  method 201
  operator 112
  sums 110
Folded functions 105, 122

Fourier series
  estimate of coefficients of 169
  Parseval equation for 97
  of functions of bounded variation 93
Fourier-Stieltjes integral 154
Fourier
  sums 90
  transform 102, 103
  transformation of integrable functions 120
Functional
  linear 46, 49
  norm of 46

Geometric mean of a function 256
Gram's determinant 16
Green's function 282, 283

Haar's problem 67
Hadamard's
  determinant theorem 18
  inequality 18
Half-line 132
Half-plane 132
Hermitian form 277
  discriminant of 274
Hölder's inequality 5

Ideal 226

Jacobi's elliptic function 62
Jensen's inequality 257

Kernels
  Fejér's 113, 118
  Krein's 177
  table of 119
Kolmogoroff 205, 261, 303, 304
Korkin 88
Krein 11, 147, 155, 254, 261, 263, 301, 303

Laplace transform 300
Legendre's polynomials 42
Levitan's polynomials 146
Limit in the mean (limes in medio) 102
Linear normalized space 2, 3, 226
Levi, Beppo
  inequality of 21
  generalized inequality of 21

Markoff, A. A. 82, 85, 269, 299, 300
  biorthogonal system of 83, 269
Markoff, W. A. 300
Minkowski's inequality 4, 7
Möbius function 268
Modulus of continuity 161, 163
Moment theorem 299
Multimonotonic functions 125

Nagy, B. Sz.-, criterion 183
Norm 2
  of a functional 46

Operator
  Fejér's 112
  Fourier's 102
  integral 113
  unitary 105
  inverse 104
Orthogonal vectors 8
Orthogonalization
  of vector systems 23
  Schmidt's process of 24
Orthogonalized vector systems 22
  complete 25, 40
  infinite 25

Parseval's equation 26
  generalized 28
  for Fourier series 97
  for pairs of functions 105
Periodic functions 64, 195
Plancherel's theory 104, 108, 134
Polya 134
Product
  inner 29
  scalar 8

Relative error 62, 64, 287
Ring, normalized 226

Schwarz's inequality 6
  generalization of 9, 17
Sets 14, 49
Slit plane 135, 280, 282
Solotareff 64, 88, 298
  problem of 280
Space
  c 7
  C 3, 49
  H 8
  l^p 7
  L 4, 255
  L^2 8, 40, 105
  L^p 4, 12, 40
  m 7
  Banach (L^p, C) 3, 4, 12, 45
  complete 15, 27
  Hilbert 8
  metric 1
  non-separable 29
  separable 14, 33
  strictly normalized 11
Stekloff function 123, 173
Step function 158
Stieltjes integral 49
Subspace 21
Suchomlinoff 297
Szasz 269, 305
Szegö 256

Tauber theorem 237
  Carleman's Tauber theorem 241
Tchebysheff's
  polynomials 58, 73
  systems of functions 73
Theorem of
  Abel 237
  Bernstein 30, 138, 206, 219, 247, 250, 252, 266, 291
  Cauchy-Hadamard 221
  Cohn 277
  de la Vallée-Poussin 52, 60
  Fejér 110
  Fejér-Riesz 152
  Fekete 288
  Fischer-Riesz 41
  Fubini 7
  Hadamard 18
  Hahn-Banach 47, 50
  Ikehara 238
  Jackson 188, 198, 199
  Jordan 95
  Lebesgue 116
  Levitan 146
  Markoff, A. A. 82
  Müntz 43, 269
  Offord 122
  Plancherel 108
  Priwaloff 210
  Riemann-Lebesgue 92, 103
  Riesz 47, 263
  Rolle 196
  Schur 276
  Watson 106, 108
  Wiener 224, 234
  Wiener-Paley 134
  Wiener-Levy 230
  Young 97
  Young-Hardy 116, 119, 180
Theta series 284
Titchmarsh 106
Transfinite diameter 288
Triangular inequality 1
Trigonometric integrals 101
True maximum 49

Vectors 8, 22, 23

Weakly decreasing function 235
Weierstrass' function 66, 165, 208

Zygmund 162, 209, 303

THEORY OF APPROXIMATION

N.I. Achieser

This excellent text, written by a mathematician who pioneered many of the modern developments in approximation theory, approaches the subject from the viewpoint of functional analysis. It offers clear, cogent treatments of classical topics such as Weierstrass' first and second theorems, and it includes generalized treatments of, among others, Hadamard's theorem on determinants, Tchebysheff's approximation by rational functions, de la Vallée-Poussin's approximation theorem and S. N. Bernstein's theorems.

The first two chapters deal with approximation problems in linear normalized spaces and the ideas of P. L. Tchebysheff. Chapter III examines the elements of harmonic analysis, and Chapter IV, integral transcendental functions of the exponential type. The last two chapters deal with the best harmonic approximation of functions, and Wiener's theorem on approximation. Professor Achieser concludes this exemplary text with an extensive section of problems and applications, including elementary extremal problems, Szego's theorem and some of its applications, the Carathéodory-Fejér problem and Soltareff's problems.

BOOKS ON MATHEMATICS

AKHIEZER, N. I., and GLAZMAN, I. M. *Theory of Linear Operators in Hilbert Space.* Translated by Merlynd Nestell.
Vol. I, $6.25

— Vol. II, *In preparation,* $6.25

BAKER, HENRY F. *Principles of Geometry:* Vol. I: *Foundations.* $5.50

— Vol. II: *Plane Geometry, Conics, Circles, Non-Euclidean Geometry.* $6.50

— Vol. III: *Solid Geometry, Quadrics, Cubic Curves in Space, Cubic Surfaces.* $6.50

— Vol. IV: *Higher Geometry.* Being illustrations of the utility of the consideration of higher space, especially of four and five dimensions. $7.50

— Vol. V: *Analytical Principles of the Theory of Curves.* $6.50

— Vol. VI: *Algebraic Surfaces.* $7.50

DEAUX, Roland. *Introduction to the Geometry of Complex Numbers.* Translated by Howard Eves. $6.50

GOODSTEIN, R. L. *Mathematical Logic.* $3.75

— *Constructive Formalism:* Essays on the Foundations of Mathematics. $3.00

GOODSTEIN, R. L. and PRIMROSE, E. J. F. *Axiomatic Projective Geometry.* $3.25

LIUSTERNIK, L., and SOBOLEV, V. *Elements of Functional Analysis.* Translated by A. E. Labarre, Jr., Herbert Izbicki, and Ward Crowley $8.50

NATANSON, I. P. *Theory of Functions of a Real Variable.* Vol. I (chapters I to IX). Translated and edited by Leo F. Boron and Edwin Hewitt. $8.50

— *Theory of Functions of a Real Variable.* Vol. II (chapters X to XVII with appendices). Translated and edited by Leo F. Boron. $8.50

RIESZ, Frigyes, and SZ.-NAGY, Béla. *Functional Analysis.* Translated by Leo F. Boron. $10.50

— *Appendix to Functional Analysis:* Extensions of Linear Transformations in Hilbert Space Which Extend Beyond This Space, by Béla Sz.-Nagy. $2.50

VAN DER WAERDEN, B. L. *Modern Algebra.* Revised edition.
Vol. I, translated by Fred Blum. $6.00
Vol. II, translated by Theodore J. Benac. $5.50

WAISMANN, Friedrich, *Introduction to Mathematical Thinking.* Translated by Theodore J. Benac. $4.50